

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2005, 01:52:59 ; Search time 174 Seconds  
(without alignments)  
555.690 Million cell updates/sec

Title: US-10-771-708-11  
Perfect score: 1354  
Sequence: 1 EQLKLVGFARFVSVADSVHR.....PPETGQNFVCKSVWYYNNE 250

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues  
Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum March 0%  
Maximum March 100%  
Listing first 45 summaries

Database : A Geneseq\_16Dec04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	138	10.2	1638	4	ABB67889 Drosophila
2	125	9.2	758	4	ABB58337 Drosophila
3	121	8.9	744	4	ABB58426 Drosophila
4	120.5	8.9	692	4	ABB70535 Drosophila
5	116	8.6	790	4	AA667230 Amino aci
6	113	8.3	833	4	ABB58185 Drosophila
7	105	7.8	812	3	AA550872 Bovine pl
8	105	7.8	812	3	AA595855 Bovine pl
9	105	7.8	812	4	AA667224 Amino aci
10	104	7.7	810	4	AA667228 Amino aci
11	102	7.5	810	2	AAW31169 Plasminog
12	101.5	7.5	790	2	AA660519 Human 'G1
13	101	7.5	480	7	ABU64288 Human ang
14	101	7.5	563	5	ABB75942 Endotheli
15	101	7.5	566	2	AA702100 A multifu
16	101	7.5	571	5	ABB75944 Angiotens
17	101	7.5	576	5	ABB75943 Angiotens
18	101	7.5	654	8	ABW83810 Human dia
19	101	7.5	672	8	ABW83809 Human dia
20	101	7.5	791	3	AAW34285 Human pla
21	101	7.5	791	3	AA550867 Human pla
22	101	7.5	791	3	AA501887 Human pla
23	101	7.5	791	3	AA595859 Human pla
24	101	7.5	791	4	AA667223 Amino aci
25	101	7.5	791	5	AAE23660 Human pla

26	101	7.5	791	6	ABG76087 Human pla
27	101	7.5	791	6	ADA08448 Mammalian
28	101	7.5	791	7	ABG75024 Glu-Plasm
29	101	7.5	791	8	ADL92182 Plasminog
30	101	7.5	791	8	ADK23654 Human pla
31	101	7.5	791	8	ADQ14615 Human pla
32	101	7.5	791	8	ADP67420 Human pla
33	101	7.5	791	8	ADQ28279 Human pla
34	101	7.5	791	8	ADS20380 Human pla
35	101	7.5	807	2	AA820013 PA mutant
36	101	7.5	810	2	AA808065 Human pla
37	101	7.5	810	2	AA813221 Human pla
38	101	7.5	810	2	AA813219 R561G hum
39	101	7.5	810	2	AA812406 R561G hum
40	101	7.5	810	2	AA813220 R561S hum
41	101	7.5	810	2	AA812938 Plasminog
42	101	7.5	810	2	AA834428 Sequence
43	101	7.5	810	2	AA502114 SEQ ID 77
44	101	7.5	810	2	AA508685 Human pla
45	101	7.5	810	3	AA582690 Human pla

ALIGNMENTS

RESULT 1  
ABB67889  
ID ABB67889 standard; protein, 1638 AA.  
XX  
AC ABB67889;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster polypeptide SEQ ID NO 30459.  
XX  
KW Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200171042-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 23-MAR-2001; 2001WO-US009231.  
XX  
PR 23-MAR-2000; 2000US-0191637P.  
PR 11-JUL-2000; 2000US-00614150.  
XX  
PA (PEKE ) PE CORP NY.  
XX  
PI Venter JC, Adams M, Li PWD, Myers EW;  
XX WPI; 2001-656860/75.  
XX DR N-PSDB; ABL11992.  
XX  
PT New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signaling and cell-cell  
PT interactions.  
XX  
PS Disclosure; SEQ ID NO 30459; 21pp + Sequence Listing; English.  
XX  
CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (AB146176-AB130511), expressed DNA  
CC sequences (AB101840-AB116175) and the encoded proteins (AB857737-  
CC AB872072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pcc\_sequences

SQ Sequence 1638 AA;

10.2%; Score 138; DB 4; Length 1638;

Query Match  
Best Local Similarity 21.2%; Pred. No. 0.00033;  
Matches 58; Conservative 43; Mismatches 100; Indels 72; Gaps 13;

QY 3 KILVGFARVVSADSVHRCISACINAFDTFGFECESVMYYPVDAECILNTEDELDLPDLF 62  
 Db 211 KLPISDIKELITANR--SDGEDKCLN---EFSFVCRSANFPMSTRSCTLSFTTRTHPELM 266  
 QY 63 VDHEDTVITLDNNKACGCECHMFHFNFKISGILINDQFAIAQCVAPYTVYAVEGROL 122  
 Db 267 EDDPNSD--YLENYTCINAE-----RRCDG--LAVFKEENKRL 300  
 QY 123 SD--ELDHSFEGELSECELCCTQRLSVTANDFNCKSFMYSNLTRSCVLSDEHS-RPLGR 179  
 Db 301 GGFPEVD-ITNNNTLBECCQTMG-----LRAEKTFGRSVERPDOSKOCILSEEDSISQKOD 354  
 QY 180 ANLAEPVPGWTYFE-----SRGVPSFTRVPQMLVG--FASF 213  
 Db 355 ISISSPTHHFYDLVCLDNGRANDYPDNSVTSHLFSGRRPDTAFQRYNRSRLGGEFHS 414  
 QY 214 VMENVPSTWCLDQCTSPPEPTQNFVCKSVMY 246  
 Db 415 ITGR--SLSECLDECLRQ-----TSFQCRSAVY 440

## RESULT 2

ABBS8337  
ID ABBS8337 standard; protein; 758 AA.

XX ABBS8337;

DT 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 1803.

XX Drosophila; developmental biology; cell signalling; insecticide;

XX pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

XX 11-JUL-2000; 2000US-00614150.

XX (PEKE ) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX MPI; 2001-656860/75.

XX N-PSDB; ABL02440.

PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions.

XX Disclosure; SEQ ID NO 1803; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-  
 CC ABB72072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 758 AA;

9.2%; Score 125; DB 4; Length 758;

Query Match  
Best Local Similarity 25.6%; Pred. No. 0.0025;  
Matches 62; Conservative 25; Mismatches 93; Indels 62; Gaps 13;

QY 21 CTSACINAFDTFGFECESVMYYPVDAECIL--NTEDRLDRPLFVDEHEDTVITLDNNKA 78  
 Db 160 CLEAC-----QANESCSAVNY--ETGLCVMFSTADQLP--GSLSSQYVPFTVYQKSCF 211  
 QY 79 G---CECHWFHFNFKISGILINDQFAIAQCVAPYTVYAVEGROLSEDLHSFEGLEL 135  
 Db 212 GVRPCKAMCITR-----VQGYRLPERAKASQSVATR 243  
 QY 136 SECELCCTQRLSVTANDFNCKSFMYSNLTRSCVLSD-EKSRPLGRANLAEPVPGWTYFESR 194  
 Db 244 RDCIEIC-----LGETEFGRSANYAHSGLCELSDMRITLSDBANIAAYDGA DYLENN 298  
 QY 195 GV--PS-----FTRVPQMLLVGFASFVMENVPSTWCLDQCTSPPEPTQNFVCKSVMY 248  
 Db 299 CAEPSTKCEFRVYAGRIKTVDS--VHONVQTLDECRDLCTAP-----FRCHS--YDY 349  
 QY 249 NE 250  
 Db 350 NE 351

## RESULT 3

ABBS8426  
ID ABBS8426 standard; protein; 744 AA.

XX ABBS8426;

DT 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 2070.

XX Drosophila; developmental biology; cell signalling; insecticide;

XX pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

XX 11-JUL-2000; 2000US-00614150.

XX (PEKE ) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX MPI; 2001-656860/75.

XX N-PSDB; ABL02529.

PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions.

XX Disclosure; SEQ ID NO 2070; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-

CC ABB72072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 744 AA;

Query Match 8.9%; Score 121; DB 4; Length 744;

Best Local Similarity 23.4%; Pred. No. 0.0061;  
Matches 64; Conservative 29; Mismatches 102; Indels 78; Gaps 14;

```
OY 5 LVGFAREVNSA-----DS-----VHRCISACINAEPTGFECESVMYYPVDAEC 48
DB 117 LVGF--EIVTGVSFAPSEKMDSGGTMLTDCIDTCRK-----NKTGQSVNY--ETGLC 167
OY 49 ILNTEDRIDRPD-LEVDHEDETVIYLDNNKAG---CECHWHFNDNKTGILNDQFALIA 104
DB 168 VLFSGHADQLPGALTKSGFPVFTIYAKRSCIAVKRCBAMWYDR----- 211
OY 105 QCVAPYVYQYVAVEGRQSLDELDHSECELESECELTQRLSVTANPFNCKSPFVSNLT 164
DB 212 -----VQNYKLKTEYKRTVSVASRRECELC-----LGENDTGRSANYDRTS 254
OY 165 RSCVLSDESRSPPLGRANLAEV-PCWTFYESRGP-----SFTRVPMILVGFASFVME 217
DB 255 GACBELSELDRILTLASGSAFQVNDGSDYLENHCVDSPNKLCEFKRLPGRILKTVDS-VYQE 313
OY 218 VPSVTMCLDQCTSPPEPTGQNFVCKSVMYTYNE 250
DB 314 VSSIDECRELCLNSP-----YRCHS--YDYND 338
```

## RESULT 4

ABB70535  
ID ABB70535 standard; protein; 692 AA.

XX AC ABB70535;

XX DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 38397.

XX KW Drosophila; developmental biology; cell signalling; insecticide;

XX KW pharmaceutical.

XX OS Drosophila melanogaster.

XX PN WO200171042-A2.

XX PD 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US009231.

XX PR 23-MAR-2000; 2000US-0191637P.

XX PR 11-JUL-2000; 2000US-00614150.

XX BA (PEKE ) PE CORP NY.

XX PI Venter JC, Adams M, Li PWD, Myers EW;

XX DR WPI; 2001-656860/75.

XX DR N-PSDB; ABL14638.

XX PT New isolated nucleic acid detection reagent for detecting 1000 or more  
XX interactions.

XX PS Disclosure; SEQ ID NO 38397; 21bp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins (AB57737-  
CC ABB72072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 692 AA;

Query Match 8.9%; Score 120.5; DB 4; Length 692;

Best Local Similarity 19.9%; Pred. No. 0.0062;  
Matches 53; Conservative 35; Mismatches 92; Indels 87; Gaps 9;

```
OY 3 KILVFAREVNSADSVHRCISACINAEPTGFECESVMYYPVDAECILNTEDRIDRPDLF 62
DB 118 KVIKGLDNALITYSTKACISACIN---ERRFVRSYEVYDNNMKCVLSDDRSSGGFV 174
OY 63 VDEHEDTVIYLDNNC---AGCECHWHFNDNKTGILNDQFALIAQCVAPYVYQYVAE 118
DB 175 QLVDAQGTDPYENLCLKPAQACKNNRSFGNSQKGVSEK-----VAQYVGL- 221
OY 119 GRQSLDELHSECELESECELTQRLSVTANPFNCKSPFVSNLTR----- 165
DB 222 -----HYTDEKELQVTSASACRLACEISECELCFLLYLOQPGSQYNCRLYLHDH 272
OY 166 -----SCVLSDBR-----SRPLGRANLAEVPCWTFYESR-----GVPSFTRVPO 204
DB 273 KTLVDGSEFTYLNHERPLIDHEPFIQ-----YENQCEKRAAGLAGASPTOT----- 318
OY 205 MLVGFASFVMEANVPSVTMCLDQCTSP 231
DB 319 -----LDKIDTLFVSLDTIEDP 335
```

## RESULT 5

AAG67230  
ID AAG67230 standard; protein; 790 AA.

XX AC AAG67230;

XX DT 13-NOV-2001 (first entry)

DE Amino acid sequence of pig plasminogen.

XX KW

XX KW Angiostatin; plasminogen; sulphydryl donor; angiogenesis; tumour;

XX KW angiogenic disease; neoplastic disease; connective tissue disorder;

XX KW rheumatoid arthritis; atherosclerosis; ocular angiogenic disease;

XX KW cerebral vascular disease; diabetes; immune disorder;

XX KW chronic inflammation; autoimmunity.

XX OS

XX OS Sus scrofa.

XX PN WO200158921-A2.

XX PD 16-AUG-2001.

XX PF 08-FEB-2001; 2001WO-US004021.

XX PR 08-FEB-2000; 2000US-00500397.

XX PA (NOUN ) UNIV NORTHWESTERN.

XX PI Soft G, Gately ST, Twardowski P;

XX PT Producing angiostatin for treating angiogenic diseases involves  
XX contacting plasminogen with plasminogen activator and sulphydryl donor  
XX simultaneously, or producing plasmin which is contacted with sulphydryl  
XX donor.

XX PS Disclosure; Page 77-80; 101pp; English.

CC The specification describes a method for generating angiotensin in vitro.  
 CC The method comprises contacting plasminogen with a sulfhydryl donor, or  
 CC culturing cells capable of producing plasminogen activator in conditioned  
 CC culture medium (CCM) and contacting the CCM with plasminogen. Angiotensin  
 CC produced by method of the invention is useful for treating animals with  
 CC angiotensin diseases. It is useful for treating animals with  
 CC such as neoplastic diseases (e.g. tumours and tumour metastasis), benign  
 CC tumours (e.g. hemangiomas, acoustic neuromas, etc), connective tissue  
 CC disorders (e.g. rheumatoid arthritis and atherosclerosis), ocular  
 CC angiogenic diseases (e.g. diabetic retinopathy, corneal graft rejection,  
 CC etc), cardiovascular diseases, cerebral vascular diseases diabetes-  
 CC associated diseases and immune disorders (e.g. chronic inflammation and  
 CC autoimmunity). The present sequence represents a plasminogen

XX Sequence 790 AA;

Query Match 8.6%; Score 116; DB 4; Length 790;

Best Local Similarity 19.5%; Pred. No. 0.021; Mismatches 100; Indels 84; Gaps 12;

Matches 55; Conservative 43; Mismatches 100; Indels 84; Gaps 12;

QY 5 LVGFARVVSADSVHRCISACLNAPDTFGFCEESVMYYPVDAECIINTEDRLDRPDLFVD 64

DB 14 LFLSLRQVARSVEEAAKC--EAET-NFICRAFGHSDQCVMAENSKTSP---IA 67

QY 65 EHHDTV---IYLDNACGCECHMFDFNFKTSGILNDQGFALAAOCYAPYTVQYAVVG 119

DB 68 RMDVVLFEKRIYLSCKTGNGVGRGTSKTSGVYICQKWSVSPHIPKYSPEKPLAG 127

QY 120 RQ-----LSDELHSPFEGLELSECELCIQ-----RLSTANDPNC 155

DB 128 LEENYCRNPNDKEKPMWCYTTDPETRPDYCDIPECEDECHMGSGHEYGKISKTMSGIHC 187

QY 156 KS-----FMTSNLTRS-CVLSDESRPLGRANLAIEVGMTYFESRGVPSF 199

DB 188 QSWGSGSPHAGVLPKGFPMKULKMYCRNPDEGPP-----W-----CF 227

QY 200 TRPQMLVGFASFVMEANPVSVMCLDQCTSPPEPGQNFVC 241

DB 228 TYDPN-----KRMFCDDP-----RCTTPPTSGPTYQC 256

RESULT 6

ABB58185

ID ABB58185 standard; protein; 833 AA.

AC ABB58185;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 1347.

KW Drosophila; developmental biology; cell signalling; insecticide;

OS pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

XX 11-JUL-2000; 2000US-00614150.

XX (PEKE ) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX N-PSDB; ABL02288.

XX New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from Drosophila and for elucidating cell signaling and cell-cell

PS interactions.

XX Disclosure; SEQ ID NO 1347; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA

CC sequences (AB101840-AB16175) and the encoded proteins (ABB57737-

CC ABB72072). The sequence data for this patent did not form part of the

CC printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX

QY 3 KLVGFARVVSADSVHRCISACLNAPDTFGFCEESVMYYPVDAECIL- 50

DB 122 ELITGY---VLSAPSKMDLTPGLTMDCLCEACOS-----NESCNAVNY--ETGLCVLF 171

QY 51 -NTEEDRLDRPDLFVDEHEDTVIYLDNACG---CECHMFDFNFKTSGILNDQGFALAAOC 106

DB 172 KTTADKLP-GSLSRQGFVFTIYAKKSLGVPRPSKAWCTIR----- 212

QY 107 YAPYTVQYAVVGRQLSDELHSPFEGLELSECELCIQRLSVTANDFNCKSPMTSNLTRS 166

DB 213 -----VQGYRLPEHVKSQGVLSRRDCLFELC-----LGEPEFICRSANYRRHSG 257

QY 167 CVLSP-EKSRPLGRANLAIEVGMTYFESRGV--PS-----FTRVQMLLVGFASFVMEANVP 219

DB 258 CELSDMDRITLSAGSVAPYVDGADYLENNCAEBSKLCFERRISGKILKTVDN--VYODIN 316

QY 220 SVTMCIDQCTSPPEPGQNFVCKSVMYYYNE 250

DB 317 TIDECRDCLNSP-----YRCHS--YDIND 339

RESULT 7

AAV50872

ID AAV50872 standard; protein; 812 AA.

AC AAV50872;

DT 24-FEB-2000 (first entry)

DE Bovine plasminogen protein fragment.

KW Plasminogen; bovine; thrombolytic agent; streptokinase; antigenic;

XX blood clot; heart attack; treatment.

XX Bos taurus.

XX WO9957251-A2.

XX 11-NOV-1999.

XX 06-MAY-1999; 99WO-US010086.

XX 06-MAY-1998; 98US-0084392P.

XX (OKLA-) OKLAHOMA MEDICAL RES FOUND.

XX Zhang XC, Lin X, Tang JUN;

XX WPI; 2000-052966/04.

XX New thrombolytic agents derived from modified humanized streptokinase,



PF 08-FEB-2001; 2001W0-US004021.  
 XX 08-FEB-2000; 2000US-00500397.  
 XX (NOUN) UNIV NORTHWESTERN.  
 PA Soff G, Gately ST, Twardowski P;  
 PI WPI; 2001-550019/61.  
 DR Producing angiostatin for treating angiogenic diseases involves  
 XX contacting plasminogen with plasminogen activator and sulfhydryl donor  
 PT simultaneously, or producing plasmin which is contacted with sulfhydryl  
 PT donor.  
 PS Disclosure; Page 80-84; 101pp; English.  
 XX  
 CC The specification describes a method for generating angiostatin in vitro.  
 CC The method comprises contacting plasminogen with a sulfhydryl donor, or  
 CC culturing cells capable of producing plasminogen activator in conditioned  
 CC culture medium (CCM) and contacting the CCM with plasminogen. Angiostatin  
 CC produced by method of the invention is useful for treating animals with  
 CC angiogenesis diseases. It is useful for treating an angiogenic disease  
 CC such as neoplastic diseases (e.g. tumours and tumour metastasis), benign  
 CC tumours (e.g. hemangiomas, acoustic neuromas, etc), connective tissue  
 CC disorders (e.g. rheumatoid arthritis and atherosclerosis), ocular  
 CC angiogenic diseases (e.g. diabetic retinopathy, corneal graft rejection,  
 CC etc), cardiovascular diseases, cerebral vascular diseases, diabetes-  
 CC associated diseases and immune disorders (e.g. chronic inflammation and  
 CC autoimmunity). The present sequence represents a plasminogen  
 CC  
 XX Sequence 812 AA;  
 SQ  
 Query Match 7.8%; Score 105; DB 4; Length 812;  
 Best Local Similarity 18.3%; Pred. No. 0.28;  
 Matches 55; Conservative 42; Mismatches 83; Indels 120; Gaps 17;  
 QY 5 LVGFARFVVSADSVHRCLSACLNAPDTFGFCEESVYVPVDAECILNTEDRDLRPLFVD 64  
 DB 40 LLSLSRKNLAGRSVEDECAAKCEETD--FVCRAPFYHSKEQCVMMAENSKVTP--VF 93  
 QY 65 EHEDTVIYLDN---NC-----AGCECH-W-----HFDNFKTSGILNQ 98  
 DB 94 RMRDVLVLEKRIYLLRECKTGNGQTYRTGTAETKSGVTQCKWMSKSPHPK-----SPB 147  
 QY 99 QPFAIAA---QCYAP-----YVTOYVAVGROLSDLDHSPFGLSECEBELCTO- 144  
 DB 148 KFPPLAGLEENYCNPNNDENGPCYTT-----DPDKRYVYCDIPCEDECMHC 195  
 QY 145 -----RLSVTANDFNCKS-----FMYSNLTRS-CVLSDESRPLGRAN 181  
 DB 196 SGENYEGKIAKTMTSGRDCQAMDSQSPHAGYIPSKFPNNKLNKNNYCRNPDGRRP----- 250  
 QY 182 LAEVPMTVFESRGVSPFRVPQMLLVGFASFVMEVNPVSTVMCLDOCTSPPEPTGQNFVC 241  
 DB 251 -----W-----CFTTDPQ-----KRWELCDIP-----RCTTPPSSGPTYQC 282  
 RESULT 10  
 AAG67228  
 ID AAG67228 standard; protein; 810 AA.  
 AC AAG67228;  
 XX  
 DT 13-NOV-2001 (first entry)  
 XX  
 DE Amino acid sequence of monkey plasminogen.  
 XX  
 KW Angiostatin; plasminogen; sulfhydryl donor; angiogenesis; tumour;  
 KW angiogenic disease; neoplastic disease; connective tissue disorder;  
 KW rheumatoid arthritis; atherosclerosis; ocular angiogenic disease;  
 KW diabetic retinopathy; corneal graft rejection; cardiovascular disease;  
 KW cerebral vascular disease; diabetes; immune disorder;

KW chronic inflammation; autoimmunity.  
 XX  
 OS Macaca mulatta.  
 XX  
 PN W0200158921-A2.  
 XX  
 PD 16-AUG-2001.  
 XX  
 PF 08-FEB-2001; 2001W0-US004021.  
 XX  
 PR 08-FEB-2000; 2000US-00500397.  
 XX (NOUN) UNIV NORTHWESTERN.  
 PA Soff G, Gately ST, Twardowski P;  
 PI WPI; 2001-550019/61.  
 DR Producing angiostatin for treating angiogenic diseases involves  
 XX contacting plasminogen with plasminogen activator and sulfhydryl donor  
 PT simultaneously, or producing plasmin which is contacted with sulfhydryl  
 PT donor.  
 PS Disclosure; Page 90-93; 101pp; English.  
 XX  
 CC The specification describes a method for generating angiostatin in vitro.  
 CC The method comprises contacting plasminogen with a sulfhydryl donor, or  
 CC culturing cells capable of producing plasminogen activator in conditioned  
 CC culture medium (CCM) and contacting the CCM with plasminogen. Angiostatin  
 CC produced by method of the invention is useful for treating animals with  
 CC angiogenesis diseases. It is useful for treating an angiogenic disease  
 CC such as neoplastic diseases (e.g. tumours and tumour metastasis), benign  
 CC tumours (e.g. hemangiomas, acoustic neuromas, etc), connective tissue  
 CC disorders (e.g. rheumatoid arthritis and atherosclerosis), ocular  
 CC angiogenic diseases (e.g. diabetic retinopathy, corneal graft rejection,  
 CC etc), cardiovascular diseases, cerebral vascular diseases, diabetes-  
 CC associated diseases and immune disorders (e.g. chronic inflammation and  
 CC autoimmunity). The present sequence represents a plasminogen  
 CC  
 XX Sequence 810 AA;  
 SQ  
 Query Match 7.7%; Score 104; DB 4; Length 810;  
 Best Local Similarity 18.1%; Pred. No. 0.36;  
 Matches 54; Conservative 38; Mismatches 89; Indels 118; Gaps 14;  
 QY 5 LVGFARFVVSADSVHRCLSACLNAPDTFGFCEESVYVPVDAECILNTEDRDLRPLFVD 64  
 DB 33 LFSITKQGLGAGSIECAKCEEBE---FTCRSQYHSKEQCVMMAE---NRKSSIVF 86  
 QY 65 EHEDTVIYLDNNCAGCECHMFHDFNFKTS-----GILNDQQAIAAQC---APVTOYV 115  
 DB 87 RMRDVLVLP-----EKVYLSECKTGNGKNYRGVMSKTRGILTCQKWSSTSPH----- 133  
 QY 116 AVEGQLSDELHSPFGL-----LSCEBELCTO- 144  
 DB 134 ---RPTSPATHBSEGLEENYCRNPNNDGQPMCYTTDPEERPDYCDIPCEDECMHCS 189  
 QY 145 -----RLSVTANDFNCKS-----FMYSNLTRS-CVLSDESRPLGRANL 182  
 DB 190 GENYDGKISKMTSGLECCQAMDSQSPHAGYIPSKFPNNKLNKNNYCRNPDGRRP----- 243  
 QY 183 AEVPGMTVFESRGVSPFRVPQMLLVGFASFVMEVNPVSTVMCLDOCTSPPEPTGQNFVC 241  
 DB 244 -----W-----CFTTDPN-----KRWELCDIP-----RCTTPPSSGPTYQC 275  
 RESULT 11  
 AAW31169  
 ID AAW31169 standard; protein; 810 AA.  
 AC AAW31169;  
 XX  
 DT 26-FEB-1998 (first entry)

XX	Plasminogen protein for production of recombinant plasminogen.
XX	
KM	Plasminogen; alpha-1-antitrypsin; AAT; argserpins; recombinant;
KM	stabilising protein.
XX	
OS	Homo sapiens.
XX	
PH	Key
FT	Misc-difference 60
FT	/note= "encoded by ACC"
FT	Misc-difference 191
FT	/note= "encoded by GAA"
FT	Misc-difference 223
FT	/note= "encoded by AAA"
FT	Misc-difference 280
FT	/note= "encoded by GCT"
FT	Misc-difference 291
FT	/note= "encoded by CTG"
FT	Misc-difference 395
FT	/note= "encoded by GSA"
FT	Misc-difference 503
FT	/note= "encoded by CAG"
FT	Misc-difference 607
FT	/note= "encoded by TCT"
FT	Misc-difference 615
FT	/note= "encoded by CAG"
FT	Misc-difference 658
FT	/note= "encoded by TTG"
FT	Misc-difference 709
FT	/note= "encoded by GCT"
XX	
PN	US5648254-A.
XX	
PD	15-JUL-1997.
XX	
PP	14-JUL-1994;
PP	94US-00275076.
XX	
PR	15-JAN-1988;
PR	88US-00144357.
PR	04-DEC-1989;
PR	89US-00445302.
PR	28-OCT-1991;
PR	91US-00785865.
XX	
PA	(ZYMO ) ZYMOGENETICS INC.
XX	
PI	Kumar AA, Mulvihill ER;
XX	
DR	WPI; 1997-372063/34.
DR	N-PSDB; AAT89686.
XX	
PT	Production of recombinant plasminogen - by co-expression with plasminogen
PT	-processing or -stabilising protein.
XX	
PS	Example 2; Fig 6A-D; 32PP; English.
XX	
CC	This is the protein plasminogen. The encoding cDNA is used in a new
CC	process for the production of plasminogen where a first DNA sequence
CC	encoding plasminogen and at least one additional DNA sequence encoding a
CC	protein that processes or stabilises the plasminogen is introduced into a
CC	eukaryotic host cell. The protein is selected from alpha-1-antitrypsin
CC	(AAT) and its variants and Argserpins. Both the DNA sequences are
CC	operably linked to transcriptional promoter and terminator sequences. The
CC	host cell is cultured under conditions that allow the DNA sequences to be
CC	expressed and the recombinant plasminogen is isolated from the host cell.
CC	Co-expression of plasminogen and the protein gives increased yields of
CC	undegraded plasminogen
XX	
5Q	Sequence 810 AA;

Query Match	7.5%;	Score 102;	DB 2;	Length 810;
Best Local Similarity	17.7%;	Pred. No. 0.57;		
Matches	52;	Conservative	45;	Mismatches 89; Indels 108; Gaps 12
QY	5	LVGFAREVVSADSVARCISACLNAPDPTGFEESVMYYVPDAECLINTEDR-----LD RP	59	

```

Db      33  LFSVTKKQLGAGSIEBCAKCEBDEE---FYCRAFOYHSKEQOCVIMANKRSSIIRMR  89
QY      60  DLFDDEHEDDTIYLIDNNCAGCSCCHHFPNEFKTSGILINDQOFLAAQCVAPYUQTQYAVAG  119
Db      90  DVLFEKK---VYLSECKTGNGKNIRGTMSKTIQGTQKMSST---SPH-----  133
QY     120  RQSLDELPHSEGLE-----LSECEBELTQ-----  144
Db     134  RPRSPATHPEBGLEENYCRNPNDPQGPWCYTTPDEKRYDCDILCECEBCHNGSGNY  193
QY     145  --RLSVTANDPNCKS-----FMYSLNRSVLSDEKSRRLGRANLAEVPG  187
Db     194  DGIKISTWGLEQCMQSDQSPHAGHYTSPSCFPNNKULKKNYCANRPER-----ELRP  243
QY     188  WTYEBSRVSPETPRVQMLLVGFASFWVENVYMSVCMJDQCSPPRETDGQNFVC  241
Db     244  W-----CFTDPN-----KAMELCDT-----KCTTPPSSSGPTQOC  275

```

RESULT 12	
AA60519	
ID	AA60519 standard; protein; 790 AA.
XX	
AC	AA60519;
XX	
DT	25-MAR-2003 (revised)
DT	22-MAR-1995 (first entry)
XX	
DE	Human 'Glu' plasminogen.
XX	
KW	Serine protease; Factor-Xa; recognition site; plasminogen; kringleg;
KW	fusion protein cleavage; protein folding; primer;
KW	polymerase chain reaction; amplification.
XX	
OS	Homo sapiens.
XX	
PN	W09418227-A2.
XX	
PD	18-AUG-1994.
XX	
PE	04-FEB-1994; 94WO-DK000054.
XX	
PR	04-FEB-1993; 93DK-00000130.
PR	05-FEB-1993; 93DK-00000139.
PR	03-DEC-1993; 93WO-GB0002492.
XX	
PA	(DENZ-) DENZYME APS.
XX	
PI	Thogersen HC, Hollet TL, Etzerodt M;
XX	
DR	WPI; 1994-279681/34.
XX	
PT	Refolding of polypeptide molecules - using a cyclic process involving
PT	denaturing and renaturing conditions to produce a correctly folded prod.
XX	
PS	Disclosure; Page 148-50; 202pp; English.
XX	
CC	cDNA encoding kringleg domains 1 and 4 of human plasminogen (full sequence
CC	given in AA60519) was PCR amplified using primers given in AA071268-71.
CC	Amplified cDNA was linked to a sequence encoding the Factor-xa cleavage
CC	site (given in AA60503), subcloned in vector pLIC1MUCH6 so that it was
CC	linked to a hexahistidine-encoding sequence and expressed in E. coli
CC	QY13. The fusion protein was purified on an Ni2+-activated NTA-agarose
CC	column. A cyclic procedure was used to obtain correctly folded
CC	recombinant protein. (Updated on 25-MAR-2003 to correct PN field.)
XX	
SQ	Sequence 790 AA;

Query Match	7.5%;	Score 101.5;	DB 2;	Length 790;
Best Local Similarity	18.0%;	Pred. No. 0.62;		
Matches	53;	Conservative	42;	Mismatches 90;
				Indels 109;
				Gaps 13;





CC The present sequence is the protein sequence of a novel recombinant  
CC endothelial cell growth inhibitor, termed angioquiescin, which comprises  
CC the signal peptide and pre-activation peptide of plasminogen (see  
CC ABB75939) in association with Kringle 1-5. Angioquiescin may also  
CC include a C-terminal tumour-targeting peptide (see ABB75943-44).  
CC Angioquiescin, nucleic acids encoding it, vectors and host cells are used  
CC in the preparation of a medicament for treating a tumour by preventing  
CC angiogenesis (claimed). In claimed methods, the peripheral and tumour-  
CC infiltrating lymphocytes, hepatocytes, epidermal cells, myocytes or other  
CC somatic cells of a patient in need of anti-angiogenic therapy are  
CC transfected with a viral (especially retroviral) vector comprising an  
CC angioquiescin nucleic acid, and the transfected cells are then expanded  
CC and implanted into the patient. Angioquiescin is useful for inhibiting  
CC cancer metastasis, and for treating angiogenesis associated conditions  
CC such as tumour growth e.g., cancer, diabetes, etc. It is also useful for  
CC treating haemangioma, solid tumours, leukaemia, psoriasis, scleroderma,  
CC arteriovenous malformations, rheosia, neovascular glaucoma, diabetic  
CC retinopathy, arthritis, diabetic neovascularisation, macular  
CC degeneration, wound healing, peptic ulcer, intestinal adhesions,  
CC atherosclerosis, fractures, keloids, vasculogenesis, haematopoiesis or  
CC cat scratch fever. The protein can also be used as a birth control agent  
CC by preventing the vascularisation required for embryo implantation. The  
CC half maximal concentration (EC50) of angioquiescin for inhibiting  
CC endothelial cell proliferation is about 50 pM, compared with 100 nM for  
CC angioestatin  
XX  
SQ Sequence 563 AA;

Query Match 7.5%; Score 101; DB 5; Length 563;  
Best Local Similarity 18.0%; Pred. No. 0.44;  
Matches 53; Conservative 42; Mismatches 90; Indels 110; Gaps 13;

OY 5 LVGFARFVNSADSVHRCISACLNAFDTGFECESMYVYVDAECLINTEDR-----LDRP 59  
DB 33 LFSVTKKQLGAGSIEBCAKCEDEE---FTCRAFQYHSKEQCIVAMENRKSIIIMR 89  
OY 60 DLFDVDEHDTYIYLDNNCAGCECHMFPNFTSGILNQQFPAIAQCYAPYTVQVAVEG 119  
DB 90 DVVLFEKK---VYLSECKTGKNGKRYGTMSKTKNGITCKWSSST---SPH----- 133  
OY 120 RQLSDELDSFEGLE-----LSCECELECTQ----- 144  
DB 134 RPRFSPATHPSGIEBENYCRPNDDPGQPMCYTTDPEKRYDYCDILECEBECMHCNGENY 193  
OY 145 --RLSVTANDPNCKS-----FMYSNLTRS--CVLSDESRRLPYGRANLAIEVP 186  
DB 194 DGKISKTSGLEGCQAMDSQSPHAGYIPSKFPPNKKYLKKNYCRNPDRELRP----- 243  
OY 187 GWTYPESRGVPSFTFVPQMLLVGFASFVMEVNPVSTWCLDQCTSPRPENQNFVC 241  
DB 244 -W-----CFTTDPN-----KRWELCDIP-----RCTTTPPSSGPTYYQC 275

RESULT 15  
AA02100  
ID AAY02100 standard; protein; 566 AA.

AC AAY02100;  
XX  
DT 16-JUL-1999 (first entry)

DE A multifunctional protein of the invention.

XX Angiostatin; endostatin; interferon; thrombospondin;  
XX interferon-inducible protein; platelet factor 4; anti-angiogenic;  
XX anti-tumor; multifunctional protein; angiogenic-mediated disease; cancer;  
XX diabetic retinopathy; macular degeneration; arthritis;  
XX tumor cell production.

OS Synthetic.  
OS Homo sapiens.  
XX  
XX  
PN WO9916889-A1.

XX  
PD 08-APR-1999.  
XX  
XX 30-SEP-1998; 98MO-US020464.  
PF  
PR 01-OCT-1997; 97US-0060609P.  
XX  
XX (SEAR ) SEARLE & CO G D.  
PA  
XX Bolanowski MA, Caparon MH, Casperson GF, Gregory SA, Klein BK;  
PI McKeown UP;  
XX  
XX WPI; 1999-255098/21.  
XX  
XX New multifunctional proteins useful for treating angiogenic-mediated  
PT diseases.

PT Claim 5; Page 95-96; 121pp; English.

CC The specification describes multifunctional proteins which comprise  
CC combinations of angiostatin, endostatin, interferon, thrombospondin,  
CC interferon-inducible protein and platelet factor 4, and have anti-  
CC angiogenic and/or anti-tumor activity. The multifunctional protein may  
CC exhibit useful properties such as having similar or greater biological  
CC activity when compared to a single factor or by having improved half-life  
CC or decreased adverse side effects, or a combination of these properties.  
CC The proteins can be used for treating an angiogenic-mediated disease,  
CC e.g., cancer, diabetic retinopathy, macular degeneration, or arthritis.  
CC They can also be used for inhibiting the production of tumor cells  
CC (characteristic of lung, breast, ovarian, prostate, pancreatic, gastric,  
CC colon, renal, bladder cancers; melanoma, hepatoma, sarcoma and lymphoma)  
CC in a patient and for inhibiting tumor growth. The present sequence  
CC represents a multifunctional protein of the invention  
XX  
SQ Sequence 566 AA;

Query Match 7.5%; Score 101; DB 2; Length 566;  
Best Local Similarity 18.0%; Pred. No. 0.45;  
Matches 53; Conservative 42; Mismatches 90; Indels 110; Gaps 13;

OY 5 LVGFARFVNSADSVHRCISACLNAFDTGFECESMYVYVDAECLINTEDR-----LDRP 59  
DB 33 LFSVTKKQLGAGSIEBCAKCEDEE---FTCRAFQYHSKEQCIVAMENRKSIIIMR 89  
OY 60 DLFDVDEHDTYIYLDNNCAGCECHMFPNFTSGILNQQFPAIAQCYAPYTVQVAVEG 119  
DB 90 DVVLFEKK---VYLSECKTGKNGKRYGTMSKTKNGITCKWSSST---SPH----- 133  
OY 120 RQLSDELDSFEGLE-----LSCECELECTQ----- 144  
DB 134 RPRFSPATHPSGIEBENYCRPNDDPGQPMCYTTDPEKRYDYCDILECEBECMHCNGENY 193  
OY 145 --RLSVTANDPNCKS-----FMYSNLTRS--CVLSDESRRLPYGRANLAIEVP 186  
DB 194 DGKISKTSGLEGCQAMDSQSPHAGYIPSKFPPNKKYLKKNYCRNPDRELRP----- 243  
OY 187 GWTYPESRGVPSFTFVPQMLLVGFASFVMEVNPVSTWCLDQCTSPRPENQNFVC 241  
DB 244 -W-----CFTTDPN-----KRWELCDIP-----RCTTTPPSSGPTYYQC 275

Search completed: March 31, 2005, 02:03:03  
Job time : 177 secs

This page blank (uspio)



```

; APPLICANT: Twardowski, Przemyslaw
; TITLE OF INVENTION: "Methods and Compositions for Generating
; TITLE OF INVENTION: Angiostatin"
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross P.C.
; STREET: 1700 Lincoln St., Suite 3500
; CITY: Denver
; STATE: CO
; COUNTRY: USA
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/991,761A
; FILING DATE:
; CLASSIFICATION: 1642
; ATTORNEY/AGENT INFORMATION:
; NAME: Crook, Wannell M.
; REGISTRATION NUMBER: 31,071
; REFERENCE/DOCKET NUMBER: 3501-16-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 790 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-991-761A-13

```

```

Query Match      8.6%; Score 116; DB 4; Length 790;
Best Local Similarity 19.5%; Pred. No. 0.00098;
Matches 55; Conservative 43; Mismatches 100; Indels 84; Gaps 12;

```

```

QY 5 LVGFARFVVSADSVHRCLSACINAFDTFGFECESVMYYPVDAECIINTEDRLDRPLFVD 64
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 14 LFLSLRQVAVARSVEBCAAKC--EAST-NFICRAFOYHSHDOOCVMAEHSKTSF--1A 67
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 65 EHEDTVYIYLDNNCAGCECHWHFDNFKTSGILNDQFALAAQCYAPVYQYVAVEG 119
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 68 RMDVVLFEKRIYLSSECKTGNGKNGRTSKTSKVICQMSVSSPHIPIKTSPEKPLAG 127
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 120 RQ-----LSDELHDSFEGLELSECELCCTQ-----RLSVTANDFNC 155
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 128 LEENYCRNPNDKRPWCYTTPETRPDYCDIPECEDECHGSGEHYEGKISKTMGIEC 187
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 156 KS-----FMTSNLTRS-CVLSDRSRPLGRANLAIEVPGWTFFESRGVPSF 199
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 188 QSWGSGSPHAGYLPKSFPRKNLKNMYCRNPDEPRP-----W-----CF 227
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 200 TRVPQMLVGFASFVMEVNPVTCMDQCTSPPEPTGQFNC 241
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 228 TTDPM-----KRWECDF-----KCTTPPTSGPTQC 256
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 3
US-09-270-767-33129
; Sequence 33129 Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0

```

```

; SEQ ID NO 33129
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; US-09-270-767-33129

```

```

Query Match      8.3%; Score 113; DB 4; Length 256;
Best Local Similarity 21.3%; Pred. No. 0.00041;
Matches 47; Conservative 28; Mismatches 82; Indels 64; Gaps 7;

```

```

QY 3 KLVGFARFVVSADSVHRCLSACINAFDTFGFECESVMYYPVDAECIINTEDRLDRPLF 62
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 47 KEIRGLDNALITYSTKEACLSACLN---ERRFVCRSVYDYNNMKCVLSDDRSGGFV 103
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 63 VDEHEDTVIYLDNNC---AGCECHWHFDNFKTSGILNDQFALAAQCYAPVYQYVAVE 118
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 104 QLVDAQTDYFENLCLKPAQACNNRSFGNSQKMGVSEK-----VAQYVGL- 150
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 119 GRQLSDELHDSFEGLELSECELCCTQRLSVTANDFNCKSFMTSNLTR----- 165
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 151 -----HYTDELQVTSSESACRLCEISEFLCGRSFYLGQPGQSQYNCRLYHLDH 201
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 166 -----SCVLSDER-----SRPLGRANLAIEVPGWTFFESR 194
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 202 KTLPDGPSTYLNHERPLIDHGEPICQ-----YFENQ 232
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 4
US-09-270-767-48346
; Sequence 48346 Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 48346
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; US-09-270-767-48346

```

```

Query Match      8.3%; Score 113; DB 4; Length 256;
Best Local Similarity 21.3%; Pred. No. 0.00041;
Matches 47; Conservative 28; Mismatches 82; Indels 64; Gaps 7;

```

```

QY 3 KLVGFARFVVSADSVHRCLSACINAFDTFGFECESVMYYPVDAECIINTEDRLDRPLF 62
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 47 KEIRGLDNALITYSTKEACLSACLN---ERRFVCRSVYDYNNMKCVLSDDRSGGFV 103
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 63 VDEHEDTVIYLDNNC---AGCECHWHFDNFKTSGILNDQFALAAQCYAPVYQYVAVE 118
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 104 QLVDAQTDYFENLCLKPAQACNNRSFGNSQKMGVSEK-----VAQYVGL- 150
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 119 GRQLSDELHDSFEGLELSECELCCTQRLSVTANDFNCKSFMTSNLTR----- 165
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 151 -----HYTDELQVTSSESACRLCEISEFLCGRSFYLGQPGQSQYNCRLYHLDH 201
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 166 -----SCVLSDER-----SRPLGRANLAIEVPGWTFFESR 194
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 202 KTLPDGPSTYLNHERPLIDHGEPICQ-----YFENQ 232
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 5
US-08-991-761A-7
; Sequence 7 Application US/08991761A
; Patent No. 6576609
; GENERAL INFORMATION:
; APPLICANT: Soft, Gerald
; APPLICANT: Gately, Stephen

```

```

; APPLICANT: Twardowski, Przemyslaw
; TITLE OF INVENTION: "Methods and Compositions for Generating
; TITLE OF INVENTION: Angiostatin"
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross P.C.
; STREET: 1700 Lincoln St., Suite 3500
; CITY: Denver
; STATE: CO
; COUNTRY: USA
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/991,761A
; FILING DATE:
; CLASSIFICATION: 1642
; ATTORNEY/AGENT INFORMATION:
; NAME: Crook, Marnell M.
; REGISTRATION NUMBER: 31,071
; REFERENCE/DOCKET NUMBER: 3501-16-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 812 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-991-761A-7

```

```

Query Match 7.8%; Score 105; DB 4; Length 812;
Best Local Similarity 18.3%; Pred. No. 0.017;
Matches 55; Conservative 42; Mismatches 83; Indels 120; Gaps 17;

```

```

QY 5 LVGFAREVVSADSVHRCISACINAFDTGFECESVMYYPVDAECILNTEDLRLDPLD 64
DB 40 LLSLRKMLAGRSVDECAKCEETD--FVCRARFQYHSKEQCVMAENSGNTF--VF 93
QY 65 EHEDTVIYLDN---NC-----ACCECH-W-----HNDNFKTSLINDQ 98
DB 94 RMRDVLVEKGIYLECKTGNGQYRGTTAETKSGVTCQKWSATSPHVPKF-----SPE 147
QY 99 QFALNA---QCYAP-----YVQYVAVEGRQLSDELDPFEGLELSECEELCTQ- 144
DB 148 KFPPLGLEENYCRNDNDENBGMCTT-----DPDKRYDCDIECECEKCMHC 195
QY 145 -----RLSVYANDFNCKS-----FMSYSLTRS-CVLSDESRPLGRAN 181
DB 196 SGENYEGKIATKMSGRDCQAMDSQSPHAGYIPSKFPNNLKMVYCRNPDGPRP----- 250
QY 182 LAEVGWTYFESRGVSPFTRVPQMLLVGFASFVMENVSVMCLDOCTSPPETGQNFVC 241
DB 251 -----W-----CFTTDPQ-----KRWEFCDIP-----RCTTPPSSGPKYQC 282

```

```

RESULT 6
US-08-991-761A-11
; Sequence 11, Application US/08991761A
; Patent No. 6576609
; GENERAL INFORMATION:
; APPLICANT: Self, Gerald
; APPLICANT: Gately, Stephen
; APPLICANT: Twardowski, Przemyslaw
; TITLE OF INVENTION: "Methods and Compositions for Generating
; TITLE OF INVENTION: Angiostatin"
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:

```

```

; ADDRESSEE: Sheridan Ross P.C.
; STREET: 1700 Lincoln St., Suite 3500
; CITY: Denver
; STATE: CO
; COUNTRY: USA
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/991,761A
; FILING DATE:
; CLASSIFICATION: 1642
; ATTORNEY/AGENT INFORMATION:
; NAME: Crook, Marnell M.
; REGISTRATION NUMBER: 31,071
; REFERENCE/DOCKET NUMBER: 3501-16-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 810 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-991-761A-11

```

```

Query Match 7.7%; Score 104; DB 4; Length 810;
Best Local Similarity 18.1%; Pred. No. 0.021;
Matches 54; Conservative 38; Mismatches 89; Indels 118; Gaps 14;

```

```

QY 5 LVGFAREVVSADSVHRCISACINAFDTGFECESVMYYPVDAECILNTEDLRLDPLD 64
DB 33 LFSITKKQLGAGSIEBCAKCEEBE--FVCRARFQYHSKEQCVMAENSGNTF--VF 86
QY 65 EHEDTVIYLDNNAKGCBEHMFDFNFKTS-----GIINDQFALNAOCY---APVITQY 115
DB 87 RMRDVLVEKGIYLECKTGNGQYRGTTAETKSGVTCQKWSATSPHVPKF-----SPE 147
QY 116 AVEGRQLSDELDPFEGLE-----EKVYLSCEKTGNGQYRGTTAETKSGVTCQKWSATSPHVPKF-----SPE 147
DB 134 -----RPTSPATHSBGLSENYCRNPDNDGQSPWCYTTDPBERPDYCDIECECEKCMHC 189
QY 145 -----RLSVYANDFNCKS-----FMSYSLTRS-CVLSDESRPLGRAN 182
DB 190 GENYDGKISKMSGLBQAMDSQSPHAGYIPSKFPNNLKMVYCRNPDGPRP----- 243
QY 183 AEPGWTYFESRGVSPFTRVPQMLLVGFASFVMENVSVMCLDOCTSPPETGQNFVC 241
DB 244 -----W-----CFTTDPN-----KRWELCDIP-----RCTTPPSSGPKYQC 275

```

```

RESULT 7
US-08-469-486-54
; Sequence 54, Application US/08469486
; Patent No. 5739281
; GENERAL INFORMATION:
; APPLICANT: Thoegeersen, Hans Christian
; APPLICANT: Holsted, Thor Ias
; APPLICANT: Etzerodt, Michael
; TITLE OF INVENTION: Improved method for the refolding of
; TITLE OF INVENTION: proteins
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA

```

ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/469,486  
FILING DATE:  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/192,060  
FILING DATE: February 4, 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Paul T. Clark  
REGISTRATION NUMBER: 30,162  
REFERENCE/DOCKET NUMBER: 06363/002001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617 542 5070  
TELEFAX: 617 542 8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 54:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 790 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-469-486-54

Query Match 7.5%; Score 101.5; DB 1; Length 790;  
Best Local Similarity 18.0%; Pred. No. 0.039; Mismatches 90; Indels 109; Gaps 13;  
Matches 53; Conservative 42;  
QY 5 LVGFARVVSADSVHRCLSACINAFPTFGFECESVMYTPVDAECILNTEDR---LDRPD 60  
DB 14 LFSVTKQKQAGASIEBCAKCEDEE---FTCRAFQYHSKEQCVMENRKSIIIRMD 70  
QY 61 LFVDHEHDVITYILDNNCAGCECHMFDPNFKTSGILNDQOPALAAQCVAPYVYQYAVEGR 120  
DB 71 VVLFERK---VYLSECKTNGKNYRGTMSTKXGKITCQKMSST---SPH-----R 114  
QY 121 QLSDELDSFEGLE-----LSECEELCTQ----- 144  
DB 115 PRFSPATHPSEGLEENYCRNPNDPQGPWCYTTDPEKRYDYCDILECEBECMHGSGENYD 174  
QY 145 -RLSVTANDFNCKS-----FMYSNLTFS-CVLSDESRPLGRANLAEPVG 187  
DB 175 GKISKTMSGLECOAMDQSOPHAGYIPSKFPKNLKKYCRNPDRLELP----- 223  
QY 188 WTYFESRGVSPFTRVPQMLVGFASFVMEVNPVTMCLDOCTSPPETGQNFVC 241  
DB 224 W-----CFTTDPN-----KRWELCDIP-----RCTTPPSSGPTQC 255

RESULT 8  
US-08-469-658-54  
Sequence 54, Application US/08469658  
Patent No. 5917018  
GENERAL INFORMATION:  
APPLICANT: Th egeren, Hans Christian  
APPLICANT: Hollet, Thor las  
APPLICANT: Eizerodt, Michael  
TITLE OF INVENTION: IMPROVED METHOD FOR THE REFOLDING OF  
TITLE OF INVENTION: PROTEINS  
NUMBER OF SEQUENCES: 58  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA

ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/469,658  
FILING DATE: June 5, 1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/192,060  
FILING DATE: February 4, 1994  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Paul T. Clark  
REGISTRATION NUMBER: 30,162  
REFERENCE/DOCKET NUMBER: 06363/002002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617 542 5070  
TELEFAX: 617 542 8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 54:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 790 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-469-658-54

Query Match 7.5%; Score 101.5; DB 2; Length 790;  
Best Local Similarity 18.0%; Pred. No. 0.039; Mismatches 90; Indels 109; Gaps 13;  
Matches 53; Conservative 42;  
QY 5 LVGFARVVSADSVHRCLSACINAFPTFGFECESVMYTPVDAECILNTEDR---LDRPD 60  
DB 14 LFSVTKQKQAGASIEBCAKCEDEE---FTCRAFQYHSKEQCVMENRKSIIIRMD 70  
QY 61 LFVDHEHDVITYILDNNCAGCECHMFDPNFKTSGILNDQOPALAAQCVAPYVYQYAVEGR 120  
DB 71 VVLFERK---VYLSECKTNGKNYRGTMSTKXGKITCQKMSST---SPH-----R 114  
QY 121 QLSDELDSFEGLE-----LSECEELCTQ----- 144  
DB 115 PRFSPATHPSEGLEENYCRNPNDPQGPWCYTTDPEKRYDYCDILECEBECMHGSGENYD 174  
QY 145 -RLSVTANDFNCKS-----FMYSNLTFS-CVLSDESRPLGRANLAEPVG 187  
DB 175 GKISKTMSGLECOAMDQSOPHAGYIPSKFPKNLKKYCRNPDRLELP----- 223  
QY 188 WTYFESRGVSPFTRVPQMLVGFASFVMEVNPVTMCLDOCTSPPETGQNFVC 241  
DB 224 W-----CFTTDPN-----KRWELCDIP-----RCTTPPSSGPTQC 255

RESULT 9  
US-08-643-219-1  
Sequence 1, Application US/08643219  
Patent No. 5801146  
GENERAL INFORMATION:  
APPLICANT: Davidson, Donald J.  
APPLICANT: NOVEL ANTIANGIOGENIC PEPTIDES  
TITLE OF INVENTION: AND METHODS FOR INHIBITING ANGIOGENESIS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Abbott Laboratories  
STREET: 100 Abbott Park Road  
CITY: Abbott Park  
STATE: IL  
COUNTRY: USA  
ZIP: 60064



SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/832,087B  
FILING DATE: 03-APR-1997  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/643,219  
FILING DATE: 06-MAY-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Casuto, Dianne  
REGISTRATION NUMBER: 40,943  
REFERENCE/DOCKET NUMBER: 5940.US.P1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 847-938-3137  
TELEFAX: 847-938-2623  
TELEX:  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 791 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE:  
DESCRIPTION: protein  
US-08-832-087B-1

Query Match 7.5%; Score 101; DB 2; Length 791;  
Best Local Similarity 18.0%; Pred. No. 0.045;  
Matches 53; Conservative 42; Mismatches 90; Indels 110; Gaps 13;

QY 5 LVGFARVVVSADSVHRCISACINAFDTFGECESVMYYPVDAICLINTEDR-----LDRP 59  
DB 14 LFSVTKKQIAGSIECAKCEDEE---FTCRAPQYHSKEQCVCYMAENRSSIIIMR 70  
QY 60 DLFDVDEHEDVTYILDNNCAGCECHWHFDFNFKTSGILNDQOFAIAQCYAPVYQYVAVEG 119  
DB 71 DVVLFPEKK---VYLSECKTGNGKNYGTMSKYNGITCQKMSST---SPH----- 114  
QY 120 RQLSDELHSPFGLE-----LSECEELCTQ----- 144  
DB 115 RRFSPATHPSGIEBENYCNPNNDPGPWCYTTDPEKRYDCDILCECEECMHGSGENY 174  
QY 145 --RLSTVANDFNCKS-----FMYSNLTRS--CVLSDESRPLGRANILAEVP 186  
DB 175 DGIKISTKMSGLECQAMDQSPHAGYIPSKFPKNLKKYCNRPDBELRP----- 224  
QY 187 GWTFESRCVPSFTRVPQMLVGFASFMENVPVMTCLDQCTSPPETGQNFVC 241  
DB 225 -W-----CFTTDPN-----KRWELCDIP-----RCTTPPSGPGTYQC 256

RESULT 12  
US-08-851-350-1  
Sequence 1, Application US/08851350  
Patent No. 6057122  
GENERAL INFORMATION:  
APPLICANT: Abbott Laboratories  
TITLE OF INVENTION: NOVEL ANTIANGIOGENIC PEPTIDES,  
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING SAME AND METHODS  
NUMBER OF SEQUENCES: 38  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Abbott Laboratories  
STREET: 100 Abbott Park Road  
CITY: Abbott Park  
STATE: IL  
COUNTRY: USA  
ZIP: 60064-3500  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/851,350  
FILING DATE: 05-MAY-1997  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Casuto, Dianne  
REGISTRATION NUMBER: 40,943  
REFERENCE/DOCKET NUMBER: 5940.US.P2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 847-938-3137  
TELEFAX: 847-938-2623  
TELEX:  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 791 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE:  
DESCRIPTION: protein  
US-08-851-350-1

Query Match 7.5%; Score 101; DB 3; Length 791;  
Best Local Similarity 18.0%; Pred. No. 0.045;  
Matches 53; Conservative 42; Mismatches 90; Indels 110; Gaps 13;

QY 5 LVGFARVVVSADSVHRCISACINAFDTFGECESVMYYPVDAICLINTEDR-----LDRP 59  
DB 14 LFSVTKKQIAGSIECAKCEDEE---FTCRAPQYHSKEQCVCYMAENRSSIIIMR 70  
QY 60 DLFDVDEHEDVTYILDNNCAGCECHWHFDFNFKTSGILNDQOFAIAQCYAPVYQYVAVEG 119  
DB 71 DVVLFPEKK---VYLSECKTGNGKNYGTMSKYNGITCQKMSST---SPH----- 114  
QY 120 RQLSDELHSPFGLE-----LSECEELCTQ----- 144  
DB 115 RRFSPATHPSGIEBENYCNPNNDPGPWCYTTDPEKRYDCDILCECEECMHGSGENY 174  
QY 145 --RLSTVANDFNCKS-----FMYSNLTRS--CVLSDESRPLGRANILAEVP 186  
DB 175 DGIKISTKMSGLECQAMDQSPHAGYIPSKFPKNLKKYCNRPDBELRP----- 224  
QY 187 GWTFESRCVPSFTRVPQMLVGFASFMENVPVMTCLDQCTSPPETGQNFVC 241  
DB 225 -W-----CFTTDPN-----KRWELCDIP-----RCTTPPSGPGTYQC 256

RESULT 13  
US-09-132-154-1  
Sequence 1, Application US/09132154  
Patent No. 6251867  
GENERAL INFORMATION:  
APPLICANT: Davidson, Donald J.  
TITLE OF INVENTION: NOVEL ANTIANGIOGENIC PEPTIDES  
TITLE OF INVENTION: AND METHODS FOR INHIBITING ANGIOGENESIS  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Abbott Laboratories  
STREET: 100 Abbott Park Road  
CITY: Abbott Park  
STATE: IL  
COUNTRY: USA  
ZIP: 60064  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/132,154  
FILING DATE:  
CLASSIFICATION:



PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/832,087  
FILING DATE: 03-APR-1997  
APPLICATION NUMBER: 08/543,219  
FILING DATE: 06-MAY-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Casuto, Dianne  
REGISTRATION NUMBER: 40,943  
REFERENCE/DOCKET NUMBER: 5940.US.P1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 847-938-3137  
TELEFAX: 847-938-2623  
TELEX:  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 791 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE:  
DESCRIPTION: protein  
US-09-132-154-1

Query Match 7.5%; Score 101; DB 3; Length 791;  
Best Local Similarity 18.0%; Pred. No. 0.045;  
Matches 53; Conservative 42; Mismatches 90; Indels 110; Gaps 13;

QY 5 LVGFARVNSADSVHRCISACLNAPDTGFECESVMYYPVDAECILNTEDR-----LDRP 59  
DB 14 LFSVTKQLGAGSIEBCAKCEDEB---FTCRAPQYHSKEQCVIMANKSIIIMR 70  
QY 60 DLVDEHEDTVIYLDNNAGCECHWFPNFKTSGILNDQOFAIAQCYAPYVTVYVAVEG 119  
DB 71 DVLVEFKK---VYLSSECTGNKNGYRGMTSKTKNGITQKKMSST---SPH----- 114  
QY 120 RQLSDELHDSFEGL------LSCEBELCTQ----- 144  
DB 115 RPRFSPATHPSEGLEENYCRPNDDPGPWCYTTDPEKRYDYCDILECEBECMHCSENY 174  
QY 145 --RUSTVANDFNCKS-----FMYSNLTRS-CVLSDERRPLGRANLAVP 186  
DB 175 DGLSKTWSGLECCAMWSQSPHAGYIPSKFPNKLKKNYCRNPRLRP----- 224

QY 187 GMTYFESRGVPSFTFRVPMQLVGFASFVMEVNPVSTMCLDOCTSPPETGQNFVC 241  
DB 225 -W-----CFTTDPN-----KRWELCDIP-----RCTTPPSSGPTGYQC 256

RESULT 14  
US-08-991-761A-6  
Sequence 6, Application US/08991761A  
Patent No. 6576609  
GENERAL INFORMATION:  
APPLICANT: Soft, Gerald  
APPLICANT: Gately, Stephen  
APPLICANT: Twardowski, Przemyslaw  
TITLE OF INVENTION: "Methods and Compositions for Generating  
TITLE OF INVENTION: Angiostatin"  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESSES:  
ADDRESSER: Sheridan Rose P.C.  
STREET: 1700 Lincoln St., Suite 3500  
CITY: Denver  
STATE: CO  
COUNTRY: USA  
ZIP: 80203  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/991,761A  
FILING DATE:

CLASSIFICATION: 1642  
ATTORNEY/AGENT INFORMATION:  
NAME: Crook, Marnell M.  
REGISTRATION NUMBER: 31,071  
REFERENCE/DOCKET NUMBER: 3501-16-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 863-9700  
TELEFAX: (303) 863-0223  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 791 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-991-761A-6

Query Match 7.5%; Score 101; DB 4; Length 791;  
Best Local Similarity 18.0%; Pred. No. 0.045;  
Matches 53; Conservative 42; Mismatches 90; Indels 110; Gaps 13;

QY 5 LVGFARVNSADSVHRCISACLNAPDTGFECESVMYYPVDAECILNTEDR-----LDRP 59  
DB 14 LFSVTKQLGAGSIEBCAKCEDEB---FTCRAPQYHSKEQCVIMANKSIIIMR 70  
QY 60 DLVDEHEDTVIYLDNNAGCECHWFPNFKTSGILNDQOFAIAQCYAPYVTVYVAVEG 119  
DB 71 DVLVEFKK---VYLSSECTGNKNGYRGMTSKTKNGITQKKMSST---SPH----- 114  
QY 120 RQLSDELHDSFEGL------LSCEBELCTQ----- 144  
DB 115 RPRFSPATHPSEGLEENYCRPNDDPGPWCYTTDPEKRYDYCDILECEBECMHCSENY 174  
QY 145 --RUSTVANDFNCKS-----FMYSNLTRS-CVLSDERRPLGRANLAVP 186  
DB 175 DGLSKTWSGLECCAMWSQSPHAGYIPSKFPNKLKKNYCRNPRLRP----- 224

QY 187 GMTYFESRGVPSFTFRVPMQLVGFASFVMEVNPVSTMCLDOCTSPPETGQNFVC 241  
DB 225 -W-----CFTTDPN-----KRWELCDIP-----RCTTPPSSGPTGYQC 256

RESULT 15  
US-08-924-287A-1  
Sequence 1, Application US/08924287A  
Patent No. 6699838  
GENERAL INFORMATION:  
APPLICANT: Abbott Laboratories  
APPLICANT: Davidson, Donald J.  
TITLE OF INVENTION: NOVEL ANTIANGIOGENIC PEPTIDES,  
TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING SAME AND METHODS FOR INHIBITING  
TITLE OF INVENTION: ANGIOGENESIS  
FILE REFERENCE: 5940.US.P3  
CURRENT APPLICATION NUMBER: US/08/924,287A  
CURRENT FILING DATE: 1997-09-05  
PRIOR APPLICATION NUMBER: US 08/851,350  
PRIOR FILING DATE: 1997-05-05  
PRIOR APPLICATION NUMBER: US 08/832,087  
PRIOR FILING DATE: 1997-04-03  
PRIOR APPLICATION NUMBER: US 08/643,219  
PRIOR FILING DATE: 1996-05-03  
NUMBER OF SEQ ID NOS: 40  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 791  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-08-924-287A-1

Query Match 7.5%; Score 101; DB 4; Length 791;  
Best Local Similarity 18.0%; Pred. No. 0.045;  
Matches 53; Conservative 42; Mismatches 90; Indels 110; Gaps 13;

```

QY 5 LVGPAREVVSADSVHRCLSACLNAPDTEGECESVMYYPVDAECILNTEDR-----LDRP 59
Db 14 LFSVTKKQLGAGSIEECACACEDEE---FTCRAFOYHSKEQCQVIMAEHRKSSIIIRWR 70
QY 60 DLFVDEHEDTVIYLDNNCAGCECHWHEDNFKTSGIINDQGFALAAOCYAPYTQYAVBEG 119
Db 71 DVVLFEKK--VYLSBCKTGNKNGRGTMSKTYNGITCQKMSST---SPH----- 114
QY 120 RQLSDELDHSFEGLE-----LSCEBELCTQ----- 144
Db 115 RPRFSPATHSEBGLSENYCNPNDNDPOGFWCYTTDPEKRYDYCDILECEHECHGSGENY 174
QY 145 --RLSVTANDFNCKS-----FMTSNLTRS-CVLSDESRPLGRANLAEPV 186
Db 175 DGKISKTMGLQCOAMDQSPHAGYIPSKFPNNKYLKKNYCRNPDRELRP----- 224
QY 187 GWTFPESRGVPSFTRVPOMLLVGFASVFMENVPSVMCLDQCTSPPETGQNFVC 241
Db 225 -W-----CFTTDPN-----KRWELCDIP-----RCTTPPSSGPTYQC 256

```

Search completed: March 31, 2005, 02:07:44  
Job time : 44 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: March 31, 2005, 02:00:10 ; Search time 140 Seconds  
(without alignments)  
591.252 Million cell updates/sec

Title: US-10-771-708-11  
Perfect score: 1354  
Sequence: 1 EQLIVGFARVVSADSVHR.....PPETGQNFVCKSVVYYNE 250

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1407402 seqs, 331100923 residues

Total number of hits satisfying chosen parameters: 1407402

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*  
5: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*  
6: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep:\*  
17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep:\*  
19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*  
20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	116	8.6	790	10	US-09-825-751A-70
2	105	7.8	812	10	US-09-825-751A-71
3	101	7.5	567	16	US-10-741-601-413
4	101	7.5	569	9	US-09-946-893-5
5	101	7.5	576	9	US-09-946-893-8
6	101	7.5	576	9	US-09-946-893-6
7	101	7.5	791	9	US-09-967-386-1
8	101	7.5	791	14	US-10-304-287-1
9	101	7.5	791	15	US-10-360-101-257
10	101	7.5	791	16	US-10-778-423-1
11	101	7.5	791	16	US-10-753-646-1
12	101	7.5	791	16	US-10-735-577-1
13	101	7.5	810	9	US-09-946-893-2

14	101	7.5	810	14	US-10-193-656-2	Sequence 2, Appl1
15	101	7.5	810	14	US-10-237-144-1	Sequence 1, Appl1
16	101	7.5	810	15	US-10-135-872B-4	Sequence 4, Appl1
17	101	7.5	810	15	US-10-450-976-2	Sequence 2, Appl1
18	101	7.5	810	15	US-10-415-012-4	Sequence 4, Appl1
19	101	7.5	810	16	US-10-741-601-409	Sequence 140, App
20	101	7.5	810	16	US-10-872-198-140	Sequence 6, Appl1
21	99	7.3	333	15	US-10-135-872B-6	Sequence 2, Appl1
22	99	7.3	466	16	US-10-753-164-2	Sequence 1, Appl1
23	99	7.3	812	9	US-09-788-142-1	Sequence 1, Appl1
24	99	7.3	812	9	US-09-761-120-1	Sequence 81, Appl1
25	99	7.3	812	9	US-09-873-676-81	Sequence 1, Appl1
26	99	7.3	812	9	US-09-335-325-1	Sequence 1, Appl1
27	99	7.3	812	14	US-10-131-241-1	Sequence 1, Appl1
28	99	7.3	812	14	US-10-127-066-1	Sequence 1, Appl1
29	99	7.3	812	15	US-10-402-364-1	Sequence 1, Appl1
30	99	7.3	812	15	US-10-401-108-1	Sequence 1, Appl1
31	98	7.2	608	13	US-10-087-192-795	Sequence 795, App
32	95	7.0	548	15	US-10-369-493-22599	Sequence 22599, A
33	93.5	6.9	741	16	US-10-408-765A-1187	Sequence 1187, Ap
34	93.5	6.9	794	11	US-09-833-245-1292	Sequence 1292, Ap
35	93.5	6.9	811	15	US-10-363-616-394	Sequence 394, App
36	93.5	6.9	830	15	US-10-353-616-393	Sequence 393, App
37	93.5	6.9	896	15	US-10-004-378A-16	Sequence 16, Appl1
38	93.5	6.9	896	15	US-10-428-275-158	Sequence 158, App
39	93.5	6.9	914	15	US-10-428-275-156	Sequence 156, App
40	93.5	6.9	915	9	US-09-909-330-34	Sequence 34, Appl1
41	93.5	6.9	915	9	US-09-909-088B-34	Sequence 34, Appl1
42	93.5	6.9	915	9	US-09-905-291A-34	Sequence 34, Appl1
43	93.5	6.9	915	9	US-09-902-853-34	Sequence 34, Appl1
44	93.5	6.9	915	9	US-09-907-824-34	Sequence 34, Appl1
45	93.5	6.9	915	9	US-09-907-841-34	Sequence 34, Appl1

ALIGNMENTS

RESULT 1  
US-09-825-751A-70  
; Sequence 70, Application US/09825751A  
; Publication No. US20030065140A1  
; GENERAL INFORMATION:  
; APPLICANT: Curagen Corporation  
; APPLICANT: Vernier, Corine A.M.  
; APPLICANT: Fernandes, Elma R.  
; APPLICANT: Taupier, Raymond J  
; APPLICANT: Quinn, Kerry E  
; APPLICANT: Spytek, Kimberly A  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Herrman, John L.  
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 15966-750  
; CURRENT APPLICATION NUMBER: US/09/825,751A  
; PRIOR FILING DATE: 2001-04-30  
; PRIOR APPLICATION NUMBER: 60/194,314  
; PRIOR FILING DATE: 2000-04-03  
; PRIOR APPLICATION NUMBER: 60/225,693  
; PRIOR FILING DATE: 2000-08-16  
; NUMBER OF SEQ ID NOS: 85  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 70  
; LENGTH: 790  
; TYPE: PRT  
; ORGANISM: Sus scrofa  
; US-09-825-751A-70  
Query Match 8.6%; Score 116; DB 10; Length 790;  
Best Local Similarity 19.5%; Pred. No. 0.0064;  
Matches 55; Conservative 43; Mismatches 100; Indels 84; Gaps 12;  
QY 5 LVGFARVVSADSVHRCISACTIAFDTRGFECESVYPPVADACILNTEDRLDRDPLFVD 64  
DB 14 LFSLRKQVARSVECAKCC--EAET-NFICRAFGYHSKDQOCVMAENSKTSP---IA 67

```
QY      65 EHDPTV-----IYLNNKAGCECHMFDFNFKTSGLINDQQCALAAQCVAFVVTOYVAVEG 119
           |||          |||||          |||||          |||||          |||||
Db      68 RMDVDVLFEKRITLSECTKGNGKNRTGTSTSKTSGVTLQCKMSVSHPHLPKYSPEKFPLAG 127
                                           :
QY      120 RQ-----LSDELDSFEGLELSECELCTQ-----RLSYANDPNC 155
           :::::          :::::          :::::          :::::          :::::
Db      128 LEENYCRNPDEKGPWCYTTPDERFDYCDIPCEDEDCMHCOSGBHEGRKISKMTSIEC 187
           :::::          :::::          :::::          :::::          :::::
QY      156 KS-----FMYSNLTS-CYLSBERRPLGRANLAEVRCWTYFSRGVPSP 199
           :::::          :::::          :::::          :::::          :::::
Db      188 QSWGSGSPAHAGVLPKFPFNKMLKNMYCNMPGGEPR-----W-                CF 227
                                           :
QY      200 TRVPGQLLVGFASFWVENSVVTMCLDOSTSPPEPTGQNFCVC 241
           :::::          :::::          :::::          :::::          :::::
Db      228 TTDPN-----KRMFCILP-----RCTLPPPTSGLTYQC 256
```

```

1      RESULT 2
2      US-09-825-751A-71
3      ; Sequence 71, Application US/09925751A
4      ; Publication No. US20030065140A1
5      ; GENERAL INFORMATION:
6      ; APPLICANT: Curagen Corporation
7      ; APPLICANT: Vernet, Corine A.M.
8      ; APPLICANT: Fernandes, Elma R.
9      ; APPLICANT: Taupier, Raymond J
10     ; APPLICANT: Quinn, Kerry E
11     ; APPLICANT: Spytek, Kimberly A
12     ; APPLICANT: Rastelli, Luca
13     ; APPLICANT: Hartman, John L
14     ; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
15     ; FILE REFERENCE: 15966-750
16     ; CURRENT APPLICATION NUMBER: US/09/825,751A
17     ; CURRENT FILING DATE: 2001-04-30
18     ; PRIOR APPLICATION NUMBER: 60/194,314
19     ; PRIOR FILING DATE: 2000-04-03
20     ; PRIOR APPLICATION NUMBER: 60/225,693
21     ; PRIOR FILING DATE: 2000-08-16
22     ; NUMBER OF SEQ ID NOS: 85
23     ; SOFTWARE: PatentIn Ver. 2.1
24     ; SEQ ID NO 71
25     ; LENGTH: 812
26     ; TYPE: PRT
27     ; ORGANISM: Bos taurus
28     ; US-09-825-751A-71

```

	Query Match	7.8%	Score 105;	DB 10;	Length 812;
	Best Local Similarity	16.3%	Pred. No. 0.096;		
	Matches	55;	Conservative	42;	Mismatches 83;
				Indels 120;	Gaps 17;
QY	5	LIVGAREVVSADSVHRCISACINAFDTGFECESSVMYYPVDAECILNTEDRIDRDLFVD	64		
		::: ::::			
Db	40	LILSRKRLARSVEDCAKCEETD---FVCRARQYHKSKEQCVMAMENSKNT---VF	93		
QY	65	EHEDTVILDN---NC-----ACGECH-W---HFNDPFTSGILNDQ	98		
		::: ::::			
Db	94	RMRVIIIEKRIYLLECKTNGGQYRGRTAETKSSVTLQOKMSATSPHPKPF-----SPE	147		
QY	99	QFAIAA---OCYAP-----YTQYVAVEGRQLSDDELDSHSEGLSELSCBELCTQ-	144		
		::: ::::			
Db	148	KFLAGLEENYCRNPNDENGSPWCYTT-----DPKRDYDIDIPCECKCMHC	195		
QY	145	---RLSTAVNDFNCKS-----FMYSNLTFS-CVLSDESRPLGRAN	181		
		::: ::::			
Db	196	SGENYEKGIATMGSGDCQAMDSDSPHAGYIIPSKFPMKLNLMKNCRNPDGHRP-----	250		
QY	182	LAEVPGMTFESRGVPSFTVRPQMLLVGAFSWMENVSVMCLDQCTSPSPETQANVC	241		
		-----W-----CFTTDPQ-----KWEFCFDP-----RCTTPPSSGPKYQC	282		

```

US-10-741-601-413
: Sequence 413, Application US/10741601
: Publication No. US20040166519A1
: GENERAL INFORMATION
: APPLICANT: CARGILL, Michele et al.
: TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
: TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: C1001500
: CURRENT APPLICATION NUMBER: US/10/741,601
: CURRENT FILING DATE: 2003-12-22
: NUMBER OF SEQ ID NOS: 26415
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 413
: LENGTH: 567
: TYPE: PRY
: ORGANISM: Homo sapiens
US-10-741-601-413

```

Query Match	7.5% ; Score 101 ; DB 16, Length 567 ;
Best Local Similarity	18.0% ; Pred. No. 0.16 ;
Matches	53 ; Conservative 42 ; Mismatches 90 ; Indels 110 ; Gaps 13 ;
QY	5 LVGFAREVSADSVYHRCISACINAFDTFGEGCESVMYVYDAECILNTEDR-----LDRP 59
Db	33 LFSVTKQLQAGISIEECACKEBEE---FTCRAFQYHSKQGCIVMAENKSSIIIMR 89
QY	60 DLFDVEHEDVTIYLNNACGCECHMFHDFNFKTSGILNDQGFALAAQCYAPVTVYAVEG 119
Db	90 DVALFEKK---VYLSECKTGNKKNYRGTMSKTNGIKGICQKKMSR---SPH----- 133
QY	120 RQSLDELIDHSFEGIE-----LSGCEELCTO----- 144
Db	134 RPRFSPATHSSEGLEENYCRNPNDPQGPWCYTTDPPEKRYDCDILCEBEECHMCSGENTY 193
QY	145 --RLSVYANDPNCKS-----FMYSNLTFRS--CVLSDERSRPGRANLAEVP 186
Db	194 DGRISKTSMLGLEQAMDSQSPHAHGYPISKRPENKLNKKNYCRNDRRLRP----- 243
QY	187 GMTYFESRGVPSFIRVYQPMLLVGFASFVMEVNEVSVTMCDDCTSPPEGTQGNFC 241
Db	244 -W-----CFTTDPN-----KRRELCDIP-----RCTTPPPSSGGTYQC 275

```

RESULT 4
US-09-946-893-5
; Sequence 5, Application US/09946893
; Patent No. US20020072494A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yihai
; TITLE OF INVENTION: Materials and methods relating to endothelial cell growth
; TITLE OF INVENTION: Inhibitors
; FILE REFERENCE: Merbourn
; CURRENT APPLICATION NUMBER: US/09/946,893
; CURRENT FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: US 60/230,893
; PRIOR FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 569
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Angiogenesis
US-09-946-893-5

```

Query Match 7.5%; Score 101; DB 9; Length 569;  
Best Local Similarity 18.0%; Pred. No. 0.16,  
Matches 53; Conservative 42; Mismatches 90; Indels 110; Gaps 13

QY 5 LVFAAEVAVSDASHKSLACLAFLTFGEGCESWMTYPDAEGLINTEDR----LDRP 59  
::: :::  
33 LFSVTAKQLDAGSIEECVAARCEDEE--LTCAAFQVHSKRQQCVIAENKSSIIIMR 89  
::: :::

[illegible]

```

QY      60  DLPLFDEHEDIVYIYLDNNCAGCECHMHDPNKSTSLNDQOFALAAQCYAVYQYVAVE3  119
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      71  DVLFEEKK--VYUSECKTNGKNYKRGTMSTKXNGITCQKMSST--SH-----  114
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      120  RQSLDELHDSFEGL-----LSCEBELCTQ-----  144
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      115  RPRSPATHPSEGLEENYCNRPNDPQGMFCYTTDPPEKRYDCDILCECECHGSGENY  174
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      145  --RLSVTANDPNCKS-----PMTSNLTRS--CVLSDEBSRPLGRANLAEPV  186
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      175  DGIKISKTMGSLGECAMDQSOPHAHYLPSTFPMNKLEKKNYCRNPDRELPR-----  224
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      187  GMTYFESRGVPSFPRVQMLLVGASIVSMNVDSVTMCDDQCSIPSPETQGVNC  241
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      225  -W-----CFTTDPN-----KRELCDIP-----KCTTPPSGSPYQC  256
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 8
US-10-304-287-1
; Sequence 1, Application US/10304287
; Publication No. US20030083234A1
; GENERAL INFORMATION:
; APPLICANT: Walsman, David M.
; APPLICANT: Kwon, MiJung
; TITLE OF INVENTION: Anti-angiogenesis methods, compositions and uses therefor
; FILE REFERENCE: ME02-001
; CURRENT APPLICATION NUMBER: US/10/304,287
; CURRENT FILING DATE: 2002-11-26
; PRIOR APPLICATION NUMBER: US 60/333,866
; PRIOR FILING DATE: 2001-11-28
; NUMBER OF SEQ. ID NOS: 8
; SOFTWARE: Microsoft Word
; SEQ. ID NO 1
; LENGTH: 791
; TYPE: PRT
; ORGANISM: mammalian
; US-10-304-287-1

```

	Query Match	7.5%	Score 101;	DB 14,	Length 791;
	Best Local Similarity	18.0%;	Pred. No.0.24;		
	Matches	53;	Conservative	42;	Mismatches 90; Indels 110; Gaps 13.
QY	5	LVGFAREVASDSVHRCISACLNAPDFTGFECESEVMYPVDACEILNLTEDR----	IDRP	59	
Dd	14	LFSTVKCOLTAGSIEECSAKCEEDEE---FCFRAFOYSKSCQQCYIMENRKSSIIIMRR		70	
QY	60	DLFVDEHDVTIYLDDNCCAGCECHMHDPNFTSGILNQOAFALAAQCVAPVTOGVANEG		119	
Dd	71	DVLVEFKK---VYLSECKTGNGKRYRGIMSKTNGITQCKMSST-----SPH-----		114	
QY	120	RQLSDLELDHSFEGLP-----LSECEELCTQ-----		144	
Dd	115	RPRESPATHEEGEEENYCRNPDDPOGPWCYTTPDKREKYCYCDLICEEBECMHCISENY		174	
QY	145	--RESTVANIDNCKS-----FMYSNLTRS--CVLSDESRGRPGRAWLAEPV		186	
Dd	175	DGKISKTMISGLECQAMDSQSPPAHGYIPSKRPINKULKKNYCNPREFLRP-----		224	
QY	187	GWTYESRGVPSFFTRVPOMLLVGFASFVMEWVPSTVMCLDOCTSPEPTGGONFVC		241	
Dd	225	-W-----CFITTDPN-----KRMLCLQIP-----RCTTPPSGSGLTYOC		256	

```

RESULT 9
US-10-360-101-257
; Sequence 257, Application US/10360101
; Publication No. US20040009550A1
; GENERAL INFORMATION:
; APPLICANT: Moll, Gert N.
; APPLICANT: Leenhouts, Cornelis J.
; TITLE OF INVENTION: Export and modification of
; FILE REFERENCE: 2183-5673

```

```

?
? CURRENT APPLICATION NUMBER: US/10/360,101
?
? CURRENT FILING DATE: 2003-02-07
?
? PRIOR APPLICATION NUMBER: EP 02077060.8
?
? PRIOR FILING DATE: 2002-05-24
?
? NUMBER OF SEQ ID NOS: 309
?
? SOFTWARE: Patentin version 3.1
?
? SEQ ID NO: 257
?
? LENGTH: 791
?
? TYPE: PR1
?
? ORGANISM: Artificial Sequence
?
? FEATURE:
?
? OTHER INFORMATION: sequence of plasminogen
US-10-360-101-257

```

	Query Match	7.5%; Score 101; DB 15; Length 791;
	Best Local Similarity	18.0%; Pred. No. 0.24;
	Matches	53; Conservative 42; Mismatches 90; Indels 110; Gaps 13;
Oy	5 LVGAREVVSADSVHRCISACLNAFDTFFGECESVMYIPVDACLINTEDR---- <td>59</td>	59
Db	14 LFSVTCKQLTGAGSLIEBCAKCEBDE--FLCRAFOYSKKEQOCIVMAENKSSIIIRMR	70
Oy	60 DLFWDEHDFTIYLIDNNCAGECECHMFDEFKTSGLINDQCPALIAOCCAPVTOYVAEG	119
Db	71 DVLFLEKK---VYLSECKTGGKNRYGTSMKTNGIKTCQKMST---SPH-----	114
Oy	120 RQLSDELIDHSFEGILE-----LSRGREELCTO-----	144
Db	115 RPRESPTATHEEGELENYCRNPNDPOGPWCYTTPDKPRDYCDILICEEBECHMCGSENY	174
Oy	145 --RLSYANDFNFKCS-----FMVSNLTRS--CVLSDESRPLGRANLAEP	186
Db	175 DGRISKTMISGLECAAMDQSPPHAHGTYIRSKFPNNKLKKNYCRANDRELRR-----	224
Oy	187 GMTYFESRGVPSFRIVPQMILLVGASVMEAVEPVMTCLDOCTSPPRPQTGNFVC	241
Db	225 W-----CFITDPN-----KRELELDIP-----RCTTPPSBSGGTTYQC	256

```

RESULT 10
US-10-778-423-1
; Sequence 1, Application US/10778423
; Publication No. US20040132664A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Henkin, Jack
; APPLICANT: Davidson, Donald J.
; TITLE OF INVENTION: ANTIANGIOGENIC POLYPEPTIDES AND METHODS
; TITLE OF INVENTION: FOR INHIBITING ANGIOGENESIS
; FILE REFERENCE: 6738. US. 02
; CURRENT APPLICATION NUMBER: US/10/778,423
; CURRENT FILING DATE: 2004-02-13
; PRIOR APPLICATION NUMBER: US/09/967,386
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US 60/236,550
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 791
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-778-423-1

```

Query Match 7.5%; Score 101; DB 16; Length 791;  
Best Local Similarity 18.0%; Pred. No. 0.24;  
Matches 53; Conservative 42; Mismatches 90; Indels 110; Gaps 13.

QY 5 LVGFAKRVVNSADSVNRCLSAFLNAFFFGEGCESVMYPVDACILNTEDR-----LDRP 59  
          :::..:::..:::..:::..:::..:::..:::..:::..:::..:::..:::..:::  
Db 14 LFSVTKKCOLAGSIEBCPAKCEDEE---FTCRAFQYHKSKEQQCVIMAEKNSSIIIMR 70  
          :::..:::..:::..:::..:::..:::..:::..:::..:::..:::..:::..:::

OY 60 DLFWDEHDHYIYLNNNCAGECCHMHFDNFRTKSILNDGOFPALAACCPARYTVTYVAVEG 119

```

Db      71  DVLFEKK---VYLSECKTGKNGYRGTMSKNGITQCKMSST-----SPH----- 114
Qy      120  RQLSDELHDSFEGL-----LSCEBELCTQ----- 144
Db      115  RPRSPATHPSBGLSENYCRNPDPDQGPWCYTTDPBKYDYCDLLECEECMHCSENY 174
Qy      145  --RLSVTANDFNCKS-----FMYSNLTRS-CVLSDERSRPLGRANLAEVP 186
Db      175  DGLSKTMSGLECCQMSQSPHAGYIPSKFPNKLLKKNYCRNPDLRLP----- 224
Qy      187  GMTYFESRGVPSFTRVPQMLLVGFASFMENVPSVTMCLDOCTSPPTGQNFVC 241
Db      225  -W-----CFTTDPN-----KRWELCDIP-----RCTTPPSSGPTTYQC 256

```

RESULT 11  
US-10-753-646-1

```

; Sequence 1, Application US/10753646
; Publication No. US20040138127A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; TITLE OF INVENTION: NOVEL ANTIANGIOGENIC PEPTIDES,
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING SAME AND METHODS FOR INHIBITING
; FILE REFERENCE: 5940 US P3
; CURRENT APPLICATION NUMBER: US/10/753,646
; CURRENT FILING DATE: 2004-01-08
; PRIOR APPLICATION NUMBER: US/08/924,287A
; PRIOR FILING DATE: 2004-01-08
; PRIOR APPLICATION NUMBER: US 08/851,350
; PRIOR FILING DATE: 1997-05-05
; PRIOR APPLICATION NUMBER: US 08/832,087
; PRIOR FILING DATE: 1997-04-03
; PRIOR APPLICATION NUMBER: US 08/643,219
; PRIOR FILING DATE: 1996-05-03
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 791
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-753-646-1

```

Query Match 7.5%; Score 101; DB 16; Length 791;  
Best Local Similarity 18.0%; Pred. No. 0.24;  
Matches 53; Conservative 42; Mismatches 90; Indels 110; Gaps 13;

```

Qy      5  LVGFARFVNSADSVHRCISACLNAPDTRGFECESVWYYPVDAECILNTEDR-----LDRP 59
Db      14  LFSVTKQGLAGSIEBCAKCEDEB---FTCRAPQYHSKQCCVIMANKSSIIIMR 70
Qy      60  DLFVDEHEDTVIYLDNNCAGCECHWFNFKTSGLINDQCPAIAQCYAPYVTVAAVEG 119
Db      71  DVLFEKK---VYLSECKTGKNGYRGTMSKNGITQCKMSST-----SPH----- 114
Qy      120  RQLSDELHDSFEGL-----LSCEBELCTQ----- 144
Db      115  RPRSPATHPSBGLSENYCRNPDPDQGPWCYTTDPBKYDYCDLLECEECMHCSENY 174
Qy      145  --RLSVTANDFNCKS-----FMYSNLTRS-CVLSDERSRPLGRANLAEVP 186
Db      175  DGLSKTMSGLECCQMSQSPHAGYIPSKFPNKLLKKNYCRNPDLRLP----- 224
Qy      187  GMTYFESRGVPSFTRVPQMLLVGFASFMENVPSVTMCLDOCTSPPTGQNFVC 241
Db      225  -W-----CFTTDPN-----KRWELCDIP-----RCTTPPSSGPTTYQC 256

```

RESULT 12  
US-10-735-577-1  
Sequence 1, Application US/10735577

```

; Publication No. US20040142897A1
; GENERAL INFORMATION:
; APPLICANT: Walteman, David M.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Tumor Growth and Metastasis
; FILE REFERENCE: ME03-009
; CURRENT APPLICATION NUMBER: US/10/735,577
; CURRENT FILING DATE: 2003-12-12
; PRIOR APPLICATION NUMBER: US 60/433,140
; PRIOR FILING DATE: 2002-12-13
; NUMBER OF SEQ ID NOS: 160
; SOFTWARE: Microsoft Word
; SEQ ID NO 1
; LENGTH: 791
; TYPE: PRT
; ORGANISM: mammalian
US-10-735-577-1

```

Query Match 7.5%; Score 101; DB 16; Length 791;  
Best Local Similarity 18.0%; Pred. No. 0.24;  
Matches 53; Conservative 42; Mismatches 90; Indels 110; Gaps 13;

```

Qy      5  LVGFARFVNSADSVHRCISACLNAPDTRGFECESVWYYPVDAECILNTEDR-----LDRP 59
Db      14  LFSVTKQGLAGSIEBCAKCEDEB---FTCRAPQYHSKQCCVIMANKSSIIIMR 70
Qy      60  DLFVDEHEDTVIYLDNNCAGCECHWFNFKTSGLINDQCPAIAQCYAPYVTVAAVEG 119
Db      71  DVLFEKK---VYLSECKTGKNGYRGTMSKNGITQCKMSST-----SPH----- 114
Qy      120  RQLSDELHDSFEGL-----LSCEBELCTQ----- 144
Db      115  RPRSPATHPSBGLSENYCRNPDPDQGPWCYTTDPBKYDYCDLLECEECMHCSENY 174
Qy      145  --RLSVTANDFNCKS-----FMYSNLTRS-CVLSDERSRPLGRANLAEVP 186
Db      175  DGLSKTMSGLECCQMSQSPHAGYIPSKFPNKLLKKNYCRNPDLRLP----- 224
Qy      187  GMTYFESRGVPSFTRVPQMLLVGFASFMENVPSVTMCLDOCTSPPTGQNFVC 241
Db      225  -W-----CFTTDPN-----KRWELCDIP-----RCTTPPSSGPTTYQC 256

```

RESULT 13  
US-09-946-893-2

```

; Sequence 2, Application US/09946893
; Patent No. US20020072494A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yihai
; TITLE OF INVENTION: Materials and methods relating to endothelial cell growth
; FILE REFERENCE: Mewburn
; CURRENT APPLICATION NUMBER: US/09/946,893
; CURRENT FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: US 60/230,893
; PRIOR FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 810
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-946-893-2

```

Query Match 7.5%; Score 101; DB 9; Length 810;  
Best Local Similarity 18.0%; Pred. No. 0.25;  
Matches 53; Conservative 42; Mismatches 90; Indels 110; Gaps 13;

```

Qy      5  LVGFARFVNSADSVHRCISACLNAPDTRGFECESVWYYPVDAECILNTEDR-----LDRP 59
Db      33  LFSVTKQGLAGSIEBCAKCEDEB---FTCRAPQYHSKQCCVIMANKSSIIIMR 89
Qy      60  DLFVDEHEDTVIYLDNNCAGCECHWFNFKTSGLINDQCPAIAQCYAPYVTVAAVEG 119

```

```

Db      90 DVLFEKK---VYLSECKTGNGKNGRTGMSKTNGITCKMSST-----SPH----- 133
Qy      120 RQLSDLDHSFEGLE-----LSECEELCTQ----- 144
Db      134 RPRFSPATHPSEGLSEENYCRNPNDPOGWCYTTDPEKRYDYCDILECEBECMHGSGENY 193
Qy      145 --RLSVTANDFNCKS-----FMYSNLTRS-CVLSDBRSRPLGRANLAEPV 186
Db      194 DGKISKTMSGLECGQAMDSQSPHAGYIPSKFPKMKLKKNYCRNPDRLELP----- 243
Qy      187 GWTYFESRGVPSFTRYPQMLVGFASFVMEVNSVTMCLDOCTSPPEPTGQNFVC 241
Db      244 -W-----CFTTDPN-----KRWELCDIP-----RCTTTPPSSGPTYQC 275

```

```

RESULT 14
US-10-193-656-2
; Sequence 2, Application US/10193656
; Publication No. US2003009673A1
; GENERAL INFORMATION:
; APPLICANT: NY, Tor
; APPLICANT: HOLMDAHL, Richard
; APPLICANT: LI, Jinan
; TITLE OF INVENTION: NOVEL DRUG TARGETS FOR ARTHRITIS
; FILE REFERENCE: 3810/10577-US3
; CURRENT APPLICATION NUMBER: US/10/193,656
; PRIOR FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US 60/304,461
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/304,490
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/305,182
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 810
; TYPE: PR
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank / P00747
; DATABASE ENTRY DATE: 1986-07-21
; RELEVANT RESIDUES: (1)..(810)
US-10-193-656-2

```

```

Query Match      7.5%; Score 101; DB 14; Length 810;
Best Local Similarity 18.0%; Pred. No. 0.25; Indels 110; Gaps 13;
Matches 53; Conservative 42; Mismatches 90;

```

```

Qy      5 LVGFAREVVSADSVHRCLSACLNAPDTFGFECESVMYYPVDAECILNTEDR-----LDRP 59
Db      33 LFSVTKKQAGAGSIECAKCEDEE---FTCRAPQYHSKEQCVIAMEBKRSIIIRMR 89
Qy      60 DLFDVDEHEDTVIYLDNNCAGCECHWFNDNFKTSIGILNDQOPALAAQCYAPVYQYVAVEG 119
Db      90 DVLFEKK---VYLSECKTGNGKNGRTGMSKTNGITCKMSST-----SPH----- 133
Qy      120 RQLSDLDHSFEGLE-----LSECEELCTQ----- 144
Db      134 RPRFSPATHPSEGLSEENYCRNPNDPOGWCYTTDPEKRYDYCDILECEBECMHGSGENY 193
Qy      145 --RLSVTANDFNCKS-----FMYSNLTRS-CVLSDBRSRPLGRANLAEPV 186
Db      194 DGKISKTMSGLECGQAMDSQSPHAGYIPSKFPKMKLKKNYCRNPDRLELP----- 243
Qy      187 GWTYFESRGVPSFTRYPQMLVGFASFVMEVNSVTMCLDOCTSPPEPTGQNFVC 241
Db      244 -W-----CFTTDPN-----KRWELCDIP-----RCTTTPPSSGPTYQC 275

```

```

RESULT 15
US-10-237-144-1
; Sequence 1, Application US/10237144

```

```

; Publication No. US20030147879A1
; GENERAL INFORMATION:
; APPLICANT: NY, Tor
; APPLICANT: LI, Jinan
; APPLICANT: HELSTROM, Sten
; APPLICANT: ERIKSSON, Per-Olof
; TITLE OF INVENTION: METHOD OF WOUND HEALING
; FILE REFERENCE: 3810/10759-US1
; CURRENT APPLICATION NUMBER: US/10/237,144
; PRIOR FILING DATE: 2002-09-05
; PRIOR APPLICATION NUMBER: US 60/317,643
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 810
; TYPE: PR
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank / 625234
; DATABASE ENTRY DATE: 2000-09-15
; RELEVANT RESIDUES: (1)..(810)
US-10-237-144-1

```

```

Query Match      7.5%; Score 101; DB 14; Length 810;
Best Local Similarity 18.0%; Pred. No. 0.25; Indels 110; Gaps 13;
Matches 53; Conservative 42; Mismatches 90;

```

```

Qy      5 LVGFAREVVSADSVHRCLSACLNAPDTFGFECESVMYYPVDAECILNTEDR-----LDRP 59
Db      33 LFSVTKKQAGAGSIECAKCEDEE---FTCRAPQYHSKEQCVIAMEBKRSIIIRMR 89
Qy      60 DLFDVDEHEDTVIYLDNNCAGCECHWFNDNFKTSIGILNDQOPALAAQCYAPVYQYVAVEG 119
Db      90 DVLFEKK---VYLSECKTGNGKNGRTGMSKTNGITCKMSST-----SPH----- 133
Qy      120 RQLSDLDHSFEGLE-----LSECEELCTQ----- 144
Db      134 RPRFSPATHPSEGLSEENYCRNPNDPOGWCYTTDPEKRYDYCDILECEBECMHGSGENY 193
Qy      145 --RLSVTANDFNCKS-----FMYSNLTRS-CVLSDBRSRPLGRANLAEPV 186
Db      194 DGKISKTMSGLECGQAMDSQSPHAGYIPSKFPKMKLKKNYCRNPDRLELP----- 243
Qy      187 GWTYFESRGVPSFTRYPQMLVGFASFVMEVNSVTMCLDOCTSPPEPTGQNFVC 241
Db      244 -W-----CFTTDPN-----KRWELCDIP-----RCTTTPPSSGPTYQC 275

```

```

Search completed: March 31, 2005, 02:10:10
Job time : 142 secs

```





R:Matthews, P.  
submitted to the EMBL Data Library, January 1996  
A:Reference number: Z19495  
A:Accession: T21967  
A:Status: preliminary; translated from GB/EMBL/DBD  
A:Molecule type: DNA  
A:Residues: 1-751 <WTL>  
A:Cross-references: UNIPROT:Q020167; EMBL:Z68342; PIDN:CAA92773.1; GSPDB:GN00022; CESP:F3  
A:Experimental source: clone F38B11  
C:Genetics:  
A:Gene: CESP:F38B11.4  
A:Map position: 4  
A:Introns: 23/1; 68/3; 116/1; 210/1; 284/1; 307/1; 359/3; 539/3; 622/2; 660/1; 695/2

Query Match 19.4%; Score 262; DB 2; Length 751;  
Best Local Similarity 27.4%; Pred. No. 5.7e-16;

Matches 73; Conservative 37; Mismatches 98; Indels 58; Gaps 8;

QY 2 QKLVGFARVVSADSVHRCLSACLNAPDTEGECESVMYYPVDAECIINTEDRLDRPDL 61  
DB 131 QSLIGHAMKVLTVDSGLSDCLSKALSQSRSDYDFLCKSAIYYETGECIMNRDKFIYPL 190  
QY 62 FVDEHSDTVI-YLDNNCAGCEC-----HHHFNFKTSGILNDQOFAIAQCYAPVYQY 114  
DB 191 FKTNIIDTLVDYFENNCAADVSCKPBEETLHW----- 220  
QY 115 VAVEGRQLSDDELHSECEELCTQRLSTVANDFNCKSFMYSLNTRSCVLSDERS 174  
DB 221 VRETEYLIDSEKQVIVESSDAQECNQLC-QNNKIGSENFPCAFASNSKQEBHLTAESS 279  
QY 175 ----RFLGRANLAEVPGWTFESRGVP-----SFTRVPMQLVGFASFVWENVPST 222  
DB 280 YVGHKDDKCFNLAFLPSGSEYFEKCYCLPTMLQCIASFEFLVANNMT--SAYKTIASLSQH 337  
QY 223 MCLDQCTSPPEPTGQNFVCKSVWY 248  
DB 338 ECLSDQCMNGAR-----CSSATYFY 357

## RESULT 3

T22486  
hypothetical protein F52B11.3 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T22486  
R:Matthews, L.  
submitted to the EMBL Data Library, November 1996  
A:Reference number: Z19569  
A:Accession: T22486  
A:Status: preliminary; translated from GB/EMBL/DBD  
A:Molecule type: DNA  
A:Residues: 1-741 <WTL>  
A:Cross-references: UNIPROT:Q9XU85; EMBL:Z682268; PIDN:CAM05199.1; GSPDB:GN00022; CESP:F5  
A:Experimental source: clone F52B11  
C:Genetics:  
A:Gene: CESP:F52B11.3  
A:Map position: 4  
A:Introns: 48/1; 298/1; 341/3; 690/2

Query Match 12.4%; Score 168; DB 2; Length 741;  
Best Local Similarity 22.5%; Pred. No. 2.5e-07;

Matches 76; Conservative 38; Mismatches 100; Indels 124; Gaps 15;

QY 5 LVGFARVVSADSVHRCLSACLNAPDTEGECESVMYYPVDAECIINTEDRLDRPDLFVD 64  
DB 137 LEGFVKKSTVVENREHCLSAKLEKE--FVCKSVFHYDTLSCELSVEDKSKRP----- 188  
QY 65 EH---EDTVIYLDNNCAG-----C--EC 82  
DB 189 TVRMSSEKIDYDNNCLSRONRCGPGGNLVFVKTTNFEIRYDHTQSVGAQESCLQXC 248  
QY 63 HHHFNFKTSGILN-----DQO-----FALMAQCYAPVYQY 114

DB 249 LDSLNTFCGSVFNPKENKNCIVSDEDTFSRADQOQGVGKDYRPICVAAADLSSTCRQ 308  
QY 115 VAVE---GRQSDDELHSECEELCTQRLSTVANDFNCKSFMYSLNTRSCVLSVD 171  
DB 309 AAFERPIGSSIEGEVVAQAQVTLISDCLSLCRNL-----NCKSINDRTASSCFL-- 359  
QY 172 ERSRPLGR-ANLAEVPGWTFESRGVPSF-----TRVPMQLVGF-- 211  
DB 360 ---YAVGQDANIKANPMSMDYEFNCESQFGSMALCTNEGIRFIYNTKEPYGALYAAR 416  
QY 212 ----SFVWENVPSTMCLDQCTSPPEPTGQNFVCKSV 245  
DB 417 FSTGSOVVENAKOISI-----TFPPTVSSD--CGTVI 447

## RESULT 4

protein C52B11.1 [imported] - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Nov-2001  
C:Accession: G89459  
R:Anonymous, The C. elegans Sequencing Consortium.  
Science 282, 2012-2018, 1998  
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology  
A:Reference number: A75000; MUID:99069613; PMID:9851916  
A>Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/Projects/C\_elec  
A:Accession: G89459  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-282 <STO>  
A:Cross-references: GB:chr\_X; PIDN:AAA82467.1; PID:91086885; GSPDB:GN00028; CESP:C52B11.1  
C:Genetics:  
A:Gene: C52B11.1  
A:Map position: X

Query Match 11.8%; Score 160; DB 2; Length 282;  
Best Local Similarity 28.9%; Pred. No. 4.1e-07;  
Matches 43; Conservative 17; Mismatches 71; Indels 18; Gaps 4;

QY 112 TVYAVBERQLSDDELHSECEELCTQRLSTVANDFNCKSFMYSLNTRSCVLSVD 171  
DB 31 THFVTDNASLQSDADPIYKATSEECISACTKURDKRDRPIVCHSFYDHSFSCITHK 90  
QY 172 ERSRPLGRANLAEVPGWTFE-----SRGV-----SFTRVPMQLVGFASFVWENVPST 221  
DB 91 EKSAPVGSQAQIENSVGKRIYFEKICLSHNPQQCAQOTFIRVDQSVLVGYA-VNMTLIDSI 149  
QY 222 TMLDQCTSPPEPTGQNFVCKSVWY 250  
DB 150 ESCAAQCV-----QBADCGSANMYFYED 171

## RESULT 5

T32444  
hypothetical protein H42K12.3 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T32444; T15821  
R:Magill, L.; Harper, M.  
submitted to the EMBL Data Library, September 1997  
A:Description: The sequence of C. elegans cosmid H42K12.  
A:Reference number: Z21169  
A:Accession: T32444  
A:Status: preliminary; translated from GB/EMBL/DBD  
A:Molecule type: DNA  
A:Residues: 1-522 <MAG>  
A:Cross-references: UNIPROT:O17347; EMBL:AF026207; PIDN:AAB71266.1; GSPDB:GN00028; CESP:H  
A:Experimental source: strain Bristol N2; clone H42K12  
R:Martin, J.  
submitted to the EMBL Data Library, November 1995  
A:Description: The sequence of C. elegans cosmid C52B11.  
A:Reference number: Z18411

[illegible]

```

RESULT 6
T29557
hypothetical protein Cl6D9.1 - Caenorhabditis elegans
C|Species: Caenorhabditis elegans
C|Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C|Accession: T29557
R|Galtung, S.; Le, T.T.
submitted to the EMBL Data Library, July 1996
A|Description: The sequence of C. elegans coemid Cl6D9.
A|Reference number: Z20640
A|Accession: T29557
A|Status: preliminary; translated from GB/EMBL/DDBJ
A|Molecule type: DNA
A|Residues: 1-342 <GAT>
A|Cross-references: UNIPROT:Q22902; EMBL:U64858; PIDDN:AAB18288.1; GSPDB:IGN00023; CESP:C
A|Experimental source: strain Bristol N2; clone Cl6D9
C|Genetics:
A|Gene: CESP:Cl6D9.1
A|Map position: 5
A|Introns: 59/2; 316/3

Query Match      11.3%; Score 153.5; DB 2; Length 342;
Best Local Similarity 38.2%; Pred. No. 2.1e-06;
Matches 29; Conservative 18; Mismatches 26; Indels 3; Gaps 2;

Db -              121 KVLVGIVDQLRVDAIVNDCAQCSNGTKYDITCKSAMYYEKXECILASQSADIPDL 180
Oy               62 FVDEHEDEVITYLDNNC 77
                  | : | : | : | : | : |
Db                181 FID--DDKSTYLENSC 194

RESULT 7
T23990
hypothetical protein R07A4.4 - Caenorhabditis elegans
C|Species: Caenorhabditis elegans
C|Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C|Accession: T23990
R|Cottage, A.
submitted to the EMBL Data Library, November 1995
A|Reference number: Z19827
A|Accession: T23990
A|Status: preliminary; translated from GB/EMBL/DDBJ

```

A: Molecule type: DNA  
A: Residues: 1-351 <Wt>  
A: Cross-references: UNIPROT:Q21782, EMBL:Z67756, PDB:CAA9164.1, GSPDB:GN00028, CESP:R07A4  
A: Experimental source: clone R07A4  
C: Genetics:  
A: Gene: CESP:R07A4.4  
A: Map position: X  
A: Introns: 24/3, 47/3, 72/3, 126/1, 147/1, 169/3, 227/1, 247/1, 293/3, 321/1

Query Match	9.3%	Score 126;	DB 2;	length 351;
Best Local Similarity	25.0%	Pred. No. 0.00072;		
Matches	45;	Conservative	31;	Mismatches 72; Indels 32; Gaps 7
Qy	95	LNDQC-----FA-----IAACQAYAYTQYAVAVGROLSDLDHSFEGLEISECEL	141	
Db	8	INQQQCFDNTLNFVAVLPIVPCDPLGSHSRVQIGIEVVEAREATITFOVALIECQVA	67	
Qy	142	CTORLSVTANDN-----CKSFMYISLUTRSCTVLSDBERRRPLGRANLAEVPCWTFEESGVP	197	
Db	68	C-RVSTADGSRLLPLCRSAHFNFNATQCVSVDANPNPGYLEYKKNQNVIIYEKICIP	125	
Qy	198	S-----STRYPOMILVCGFASPVVNEAVPSVTMLCDQCTSPPEPTGONFCKSVMY	247	
Db	126	DYVLPMSCDQVRRRIPOHILIGHASEVI-SVASENECVLECIR-AKTLASVACHSTLHY	182	

```

RESULT 8
T25804
hypothetical protein H03E18.1 - Caenorhabditis elegans
C|Species: Caenorhabditis elegans
C|Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C|Accession: T25804
R|Bentley, D.
submitted to the EMBL Data Library, August 1996
A|Description: The sequence of C. elegans cosmid H03E18.
A|Reference number: Z20090
A|Accession: T25804
A|Status: preliminary; translated from GB/EMBL/DBJ
A|Molecule type: DNA
A|Residues: 1-1127 <BEN>
A|Cross-references: UNIPROT:Q94248; EMBL:U67947; PIDD:AA07557.1; GSPDB:GN00028; CESP:H03E18
A|Experimental source: strain Bristol N2; clone H03E18
C|Genetics:
A|Gene: CESP:H03E18.1
A|Map position: X
A|Intons: 112/3; 217/1; 282/2; 304/1; 424/1; 461/2; 579/3; 615/2; 831/1; 999/2; 1088/1

```

[illegible]

RESULT 9  
PLPG  
plasmin (EC 3.4.21.7) precursor - pig (fragment  
N;Alternate names: plasminogen  
N;Contains: miniplasminogen

C:Species: Sus scrofa domestica (domestic pig)  
 C:Date: 07-Sep-1990 #sequence revision 01-Nov-1996 #text\_change 09-Jul-2004  
 C:Accession: S03733; S03737; A25834  
 R:Schaller, J.; Marti, T.; Roesseler, S.J.; Kamper, U.; Rickli, E.E.  
 R:Rinhold, J., 91-102, 1987  
 A:Title: Amino acid sequence of the heavy chain of porcine plasmin. Comparison of the ca  
 A:Reference number: S03733  
 A:Accession: S03733  
 A:Molecule type: protein  
 A:Residues: 1-560 <SCH>  
 A:Cross-references: UNIPROT:P06867  
 R:Brumsholtz, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Lergier, W.; Manneberg, M.;  
 Eur. J. Biochem. 114, 465-470, 1981  
 A:Title: Comparison of the primary structure of the N-terminal CNBR fragments of human,  
 A:Reference number: S03735; PMID:8121097; PMID:7238497  
 A:Accession: S03737  
 A:Molecule type: protein  
 A:Residues: 1-57 <BRU>  
 R:Marli, T.; Schaller, J.; Rickli, E.E.  
 Eur. J. Biochem. 149, 279-285, 1985  
 A:Title: Determination of the complete amino-acid sequence of porcine miniplasminogen.  
 A:Reference number: A25834; PMID:85203907; PMID:3846533  
 A:Accession: A25834  
 A:Molecule type: protein  
 A:Residues: 450-790 <MAR>  
 C:Function:  
 A:Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a va  
 as the walls of the graafian follicle; also activates the urokinase-type plasminogen act  
 A:Pathway: fibrinolysis  
 C:Superfamily: plasmin; kringe homology; plasminogen-related protein precursor homology  
 C:Keywords: fibrinolysis; glycoprotein; hydrolase; kidney; kringe; plasma; serine prote  
 F:1-790/Product: plasminogen #status predicted <PRO>  
 F:1-77/Domains: plasminogen-related protein precursor homology (fragment) <PLPH>  
 F:1-77/Domains: activation peptide #status predicted <APT>  
 F:18-560/Product: plasmin chain A #status predicted <ACH>  
 F:84-162/Domains: kringe homology <KR1>  
 F:166-243/Domains: kringe homology <KR2>  
 F:256-333/Domains: kringe homology <KR3>  
 F:358-435/Domains: kringe homology <KR4>  
 F:450-790/Product: miniplasminogen #status experimental <MIN>  
 F:461-540/Domains: kringe homology <KR5>  
 F:561-790/Product: plasmin chain B #status experimental <BCH>  
 F:561-783/Domains: trypsin homology <TRY>  
 F:50-54,34-42,84-162,105-145,133-157,166-243,169-297,187-226,215-238,256-333,277-316,305  
 bonds: #status predicted  
 F:602,645,740/Active site: His, Asp, Ser #status predicted

Query Match 8.6%; Score 116; DB 1; Length 790;  
 Best Local Similarity 19.5%; Pred. No. 0.016;  
 Matches 55; Conservative 43; Mismatches 100; Indels 84; Gaps 12;

5 LVGFAREVVSADSVHRCISACINAFDTPGFECESVWYVVDACILNTEDRDLRPLFVD 64  
 14 LLSLSKQVLAASVECAKAC--EART-NFICRAFYHKSDDQCVMAENKSTSP---IA 67  
 65 EHEDTV-----IYLDNNGACGECHEMFNPKTSGILINDQAFIAACVAPYVTVYAVEG 119  
 68 RMRDVLLEFKRIVLYSECKTGNGKNYGTSTKISGVICQKMSVSPHPIKYSPEKPLAG 127  
 120 RQ-----LSDELHSPFELLESSECELTQ-----RLSVTANDPNC 155  
 128 LEENYCRNDNDEKGFWCYTTPTFRFDVCDIPECEDECMHSGEHYEGKISKTSGLIEC 187  
 156 KS-----FMYSNLTSS--CVLSDESRSPPLGRANLAIEVPGWTYFESRGVPSF 199  
 188 QSMGSGSPRAHAGVLPSPFNKILKNYCNPNPGEPRP-----W-----CF 227  
 200 TRVPQMLVGFASFWENVPVVTMCIDQCTSPPTGQNFVC 241  
 228 TTDPN-----KRMFECIDP-----RCTTPTPTSGPTGQC 256

RESULT 10

PLBO  
 N:Plasmin (EC 3.4.21.7) precursor - bovine  
 N:Alternate names: plasminogen  
 C:Species: Bos primigenius taurus (cattle)  
 C:Date: 30-Sep-1987 #sequence revision 28-Apr-1995 #text\_change 09-Jul-2004  
 C:Accession: S45046; A25835; I45961; S03736  
 R:Berghlund, L.; Andersen, M.D.; Petersen, T.E.  
 submitted to the EMBL Data Library, May 1994  
 A:Description: Cloning and characterization of the bovine plasminogen cDNA.  
 A:Reference number: S45046  
 A:Accession: S45046  
 A:Molecule type: mRNA  
 A:Residues: 1-812 <BER>  
 A:Cross-references: UNIPROT:P06868; EMBL:X79402; NID:g949462; PIDN:CA55939.1; PID:g49496  
 A:Experimental source: liver  
 A:Note: it is uncertain whether Met-1 or Met-8 is the initiator  
 R:Schaller, J.; Moser, P.W.; Danneberg-Muller, G.A.K.; Roesseler, S.J.; Kamper, U.; Rick  
 Eur. J. Biochem. 149, 267-278, 1985  
 A:Title: Complete amino acid sequence of bovine plasminogen. Comparison with human plasmi  
 A:Reference number: A25835; PMID:85203906; PMID:3846532  
 A:Accession: A25835  
 A:Molecule type: protein  
 A:Residues: 27-334,'D',336-515,'H',517-554,'L',556-812 <SCH>  
 R:Malinowski, D.P.; Sadler, J.E.; Davie, E.W.  
 Biochemistry 23, 4243-4250, 1984  
 A:Title: Characterization of a complementary deoxyribonucleic acid coding for human and  
 A:Reference number: I45961; PMID:8502311; PMID:6148961  
 A:Accession: I45961  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 706-743,'R',745-812 <MAL>  
 A:Cross-references: GB:K02935; NID:G163551; PIDN:AAA30714.1; PID:G163552  
 R:Brumsholtz, R.A.; Lerch, P.G.; Schaller, J.; Rickli, E.E.; Lergier, W.; Manneberg, M.;  
 Eur. J. Biochem. 114, 465-470, 1981  
 A:Title: Comparison of the primary structure of the N-terminal CNBR fragments of human, t  
 A:Reference number: S03735; PMID:8121097; PMID:7238497  
 A:Accession: S03736  
 A:Molecule type: protein  
 A:Residues: 27-83 <BRU>  
 C:Function:  
 A:Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a va  
 as the walls of the graafian follicle; also activates the urokinase-type plasminogen acti  
 A:Pathway: fibrinolysis  
 C:Superfamily: plasmin; kringe homology; plasminogen-related protein precursor homology;  
 C:Keywords: duplication; fibrinolysis; glycoprotein; hydrolase; kidney; kringe; plasma;  
 F:1-26/Domains: signal sequence #status predicted <SIG>  
 F:8-103/Domains: plasminogen-related protein precursor homology <PLPH>  
 F:27-812/Product: plasminogen #status experimental <PRO>  
 F:27-103/Domains: activation peptide #status experimental <APT>  
 F:104-583,584-812/Product: plasmin #status experimental <MAT>  
 F:110-188/Domains: plasmin chain A #status experimental <ACH>  
 F:192-269/Domains: kringe homology <KR1>  
 F:282-359/Domains: kringe homology <KR2>  
 F:384-461/Domains: kringe homology <KR3>  
 F:485-564/Domains: kringe homology <KR4>  
 F:584-812/Domains: plasmin chain B #status experimental <BCH>  
 F:584-805/Domains: trypsin homology <TRY>  
 F:56-80,60-68,110-188,131-171,159-183,192-269,195-323,213-252,241-264,282-359,303-342,331  
 bonds: #status predicted  
 F:315/Binding site: carbohydrate (Asn) (covalent) #status experimental  
 F:365/Binding site: carbohydrate (Ser) (covalent) #status experimental  
 F:64,66,762/Active site: His, Asp, Ser #status predicted

Query Match 7.8%; Score 105; DB 1; Length 812;  
 Best Local Similarity 18.3%; Pred. No. 0.017;  
 Matches 55; Conservative 42; Mismatches 83; Indels 120; Gaps 17;

5 LVGFAREVVSADSVHRCISACINAFDTPGFECESVWYVVDACILNTEDRDLRPLFVD 64  
 40 LLSLSKQVLAASVECAKAC--EART-NFICRAFYHKSDDQCVMAENKSTSP---VF 93  
 65 EHEDTVLYLDN-----NC-----AGCECH-W-----HFDNFKTSGILINDQ 98

```

Db      94 RMRDYLVEKRIYLLCEKTCGNQYRGTTAETKSGVTCQKMSATSPHYPKF-----SPE 147
      99 QFATAA-----QCYAP-----YVTVAVNEGQSLDELHDSFEGJLSECELTQ- 144
      148 KFLPLGLEENYCRNDNDENGCMYCTT-----DPRKRDYCDIDCECECKMHC 195
      145 -----RLSYTANDFNCKS-----FMYSNLTFRS-CVLSDESRPLGRAN 181
      196 GSENVGKIATMSGRDCQAMDOSPHAHGYTPSKFPRKNLKMVCRNPDEPRP----- 250
      182 LAEVPGMTYFESRGVPSFTRVPQMLLVGFASFVMEVNPVSVTMCLDOCTSPPEPTGQNFVC 241
      251 -----W-----CFTTDPQ-----KRWFCIDP-----RCTTPPSSSGPTQC 282

```

# RESULT 11

```

B30848
Plasmin (EC 3.4.21.7) precursor - rhesus macaque
C:Species: Macaca mulatta (rhesus macaque)
C>Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 09-Jul-2004
C:Accession: B32866; B30848
R:Tomlinson, J.E.; McLean, J.W.; Lawn, R.M.
J. Biol. Chem. 264, 5957-5965, 1989
A:Title: Rhesus monkey apolipoprotein(a). Sequence, evolution, and sites of synthesis.
A:Reference number: A32869; MUID:89174660; PMID:2925643
A:Accession: B32869
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-810 <TOM>
A:Cross-references: UNIPROT:P12545; GB:J04697; NID:9342272; PIDN:AAA6901.1; PID:9342273
C:Superfamily: Plasmin; kringle homology; plasminogen-related protein precursor homology
C:Keywords: fibrinolysis; glycoprotein; hydrolyase; kringle; serine proteinase
F:1-96/Domain: plasminogen-related protein precursor homology <PIPH>
F:1-9/Domain: signal sequence #status predicted <SIG>
F:103-181/Domain: kringle homology <KR1>
F:185-262/Domain: kringle homology <KR2>
F:275-352/Domain: kringle homology <KR3>
F:377-454/Domain: kringle homology <KR4>
F:481-560/Domain: kringle homology <KR5>
F:581-803/Domain: trypsin homology <TRY>
F:49-73,53-61,103-181,124-164,152-176,185-262,188-316,206-245,234-257,275-352,296-335,32
bonds: #status predicted
F:622,665,760/Active site: His, Asp, Ser #status predicted

```

```

Query Match      7.7%; Score 104; DB 2; Length 810;
Best Local Similarity 18.1%; Pred. No. 0.21;
Matches 54; Conservative 38; Mismatches 89; Indels 118; Gaps 14;

```

```

      5 LVGFARFVVSADSVNRCLSAFLNADFTGFECESVMYYPVDAECTINTEDRLDRPDLFVD 64
      33 LFSITTKQLGAGSIECAKCEBEEB---FTCRSQYHSKQOCVIAE---NRKSLVIF 86
      65 EHEBTVIYLDNNKAGCECHMFDFNFKTS-----GILNDQFALAAQCY---APVVTQYV 115
      87 RMRDVLVE-----EKVYVLSCKTGNKNGKNGTSMKTGTITCQKMSSTSPH----- 133
      116 AVEGRQLSDELHDSFEGJL-----LSECEBELCTQ- 144
      134 -----RPTSPATHSEEGLEENYCRNPNDGQPMWCYTTPDEPRPDYCDIPCECEBCHMCS 189
      145 -----RLSYTANDFNCKS-----FMYSNLTFRS-CVLSDESRPLGRANL 182
      190 GENVDGKISKMTSGIECCAMDSQSPHAGYTPSKFPRKNLKMVCRNPDEGRPR----- 243
      183 AEVPGMTYFESRGVPSFTRVPQMLLVGFASFVMEVNPVSVTMCLDOCTSPPEPTGQNFVC 241
      244 -----W-----CFTTDPN-----KRWELCIDP-----RCTTPPSSSGPTQC 275

```

RESULT 12  
T16417  
hypothetical protein F52C9.5 - Caenorhabditis elegans

```

C:Species: Caenorhabditis elegans
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: T16417
R:Ravellio, T.
submitted to the EMBL Data Library, November 1995
A:Description: The sequence of C. elegans cosmid F52C9.
A:Reference number: 218511
A:Accession: T16417
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-498 <FAV>
A:Cross-references: UNIPROT:Q10125; EMBL:U39850; NID:g1055052; PID:g1055053; PIDN:AAA810;
A:Gene: CESP:F52C9.5
A:Introns: 45/2; 87/3; 132/1; 185/3; 207/1; 294/3; 337/1; 381/3; 430/1
C:Superfamily: Caenorhabditis elegans hypothetical protein F52C9.5

```

```

Query Match      7.6%; Score 103; DB 2; Length 498;
Best Local Similarity 19.9%; Pred. No. 0.14;
Matches 53; Conservative 27; Mismatches 65; Indels 122; Gaps 11;

```

```

      14 SADSVNRCLSAFLNADFTG-FECESVMYYPVDAECTL----- 50
      154 SSTGLHCKSHCLNS--QIGVSCRSFVYDVMNRVCDLFAHVGDQAPARLKPQTODEDS 211
      51 -----NTEDRDLRDPDLFVDEHEDTVIYLDNNKAGCECHMFDFNFKTSGL 95
      212 PPSPPSAPVIALATNTDKR-----DEHEBTEHTIED----- 243
      96 NDQOFALAA-----QCYAPVTVQYVAVNEGQSLDELHDSFEGJLSECELTORLSVTAN 151
      244 ----ITVASPSDSCPRGKQSTFLRTGQFELFSHDDQELVGVDAECAKACIEN----- 293
      152 DFNKCSFMYSNLTFRS-CVLSDESRPLGRANLAEVPGMTYFESRGVPS--FTRVPQMLLVG 209
      294 ----KKSPPY-----RSC-----PSYFFSRHPQMILVG 316
      210 FASFVMEVNPVSVTMCLDOCTSPPEPTG 236
      317 FASVSQDS-PSFEHCFTCLNKTDATG 342

```

# RESULT 13

```

T19553
hypothetical protein C29B6.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T19553
R:Dobson, R.
submitted to the EMBL Data Library, May 1996
A:Reference number: Z19141
A:Accession: T19553
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-942 <WTL>
A:Cross-references: UNIPROT:Q18298; EMBL:Z72504; PIDN:CAA96604.1; GSPDB:GN00022; CESP:C2;
A:Experimental source: clone C29B6
A:Gene: CESP:C29B6.4
A:Map position: 4
A:Introns: 21/3; 67/3; 145/2; 175/1; 243/3; 308/1; 415/2; 540/1; 616/3; 659/3; 708/2; 83;

```

```

Query Match      7.6%; Score 102.5; DB 2; Length 942;
Best Local Similarity 22.1%; Pred. No. 0.35;
Matches 36; Conservative 30; Mismatches 72; Indels 25; Gaps 6;

```

```

      13 VSADSVNRCLSAFLNADFTGFECESVMYYPVDAECTINTEDRLDRPDLFVDEHEDTVIY 72
      306 ITNGSTKCLQKCTDMEN-----CKSVYVNVLTFRSLSTSKTSKISATV-SDENYD---F 356
      73 LDNNKAGCECHMFDFNFKTSGLNDQFALAAQCYAPVTVQYVAV-GRQLSDELHDSFE 131
      357 YERSC-----PYIPTSPSSSFP1PHPTLYPEVTVIRYKLMERGRQLSSBFVNKTN 407

```

QY 132 GLESLCEBLCIORLSTVANDPCKSFMYSNLTRGVLSIDERS 174  
 DB 408 VENVODPCMSLCVN-----SKVHCELLISFSTTSNQLSTLKS 444

RESULT 14

PLHM (EC 3.4.21.7) precursor [validated] - human  
 N:Alternate names: plasminogen precursor [misnomer]  
 N:Contents: angiotatin; macroplasmin; plasminogen  
 C:Species: Homo sapiens (man)  
 C:Date: 24-Apr-1984 #sequence revision 02-Dec-1994 #text change 09-Jul-2004  
 C:Accession: A35229; I52242; A26646; I62733; I84609; S03735; A00929; A04627; A04625; A04  
 R:Peterson, T.E.; Martzen, W.R.; Ichinose, A.; Davie, E.W.  
 J. Biol. Chem. 265, 6104-6111, 1990  
 A:Title: Characterization of the gene for human plasminogen, a key proenzyme in the fibr  
 A:Reference number: A35229; MUID:90202879; PMID:2318848  
 A:Accession: A35229  
 A:Molecule type: DNA  
 A:Residues: 1-810 <P>  
 A:Cross-references: UNIPROT:P00747; UNIPROT:Q9UBQ9; UNIPROT:Q9UMI2; GB:J05286; GB:M34276  
 A:Experimental source: leukocyte; lung fibroblast  
 R:Malgaroli, N.; Bruno, L.; Pontoglio, M.; Candiani, G.; Meroni, G.; Ottolenghi, S.; Ta  
 Biochem. Biophys. Res. Commun. 173, 1013-1018, 1990  
 A:Title: Definition of the transcription initiation site of human plasminogen gene in li  
 A:Reference number: I52242; MUID:91097523; PMID:2268308  
 A:Accession: I52242  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-16 <MAL1>  
 A:Cross-references: GB:M62890; NID:9190092; PIDN:AAA3454.1; PID:9553613  
 R:Rorsgren, M.; Raden, B.; Israelsson, M.; Larsson, K.; Heden, L.O.  
 FEBS Lett. 213, 254-260, 1987  
 A:Title: Molecular cloning and characterization of a full-length cDNA clone for human pl  
 A:Reference number: A26646; MUID:87162490; PMID:3030813  
 A:Accession: A26646  
 A:Molecule type: mRNA  
 A:Residues: 1-471, 'D', 473-810 <FOR>  
 A:Cross-references: GB:X05199; NID:935530; PIDN:CAA2863.1; PID:935531  
 A:Experimental source: liver  
 R:Malinowski, D.P.; Sadler, J.E.; Davie, E.W.  
 Biochemistry 23, 4243-4250, 1984  
 A:Title: Characterization of a complementary deoxyribonucleic acid coding for human and  
 A:Reference number: I45961; MUID:85023311; PMID:6148961  
 A:Accession: I62738  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 292-471, 'D', 473-810 <MAL2>  
 A:Cross-references: GB:X02922; NID:9190112; PIDN:AAA60124.1; PID:9387031  
 A:Accession: I84609  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 367-419 <MAL3>  
 A:Cross-references: GB:X02921; NID:9190110; PIDN:AAA60123.1; PID:9190111  
 R:Brumsholtz, R.A.; Lerch, P.G.; Schaller, J.; Ricelli, E.E.; Lergier, W.; Manneberg, M.;  
 Eur. J. Biochem. 114, 465-470, 1981  
 A:Title: Comparison of the primary structure of the N-terminal CNBr fragments of human,  
 A:Reference number: S03735; MUID:81212097; PMID:7228497  
 A:Accession: S03735  
 A:Molecule type: protein  
 A:Residues: 20-71, 'E', 73-76 <BRU>  
 R:Scott-Jensen, L.; Petersen, T.E.; Magnusson, S.  
 submitted to the Atlas, July 1977  
 A:Reference number: A00929  
 A:Accession: A00929  
 A:Molecule type: protein  
 A:Residues: 20-71, 'E', 73-85, 87-106, 'D', 108-360, 'E', 362-810 <SOT>  
 R:Wiman, B.  
 Eur. J. Biochem. 76, 129-137, 1977  
 A:Title: Primary structure of the B-chain of human plasmin.  
 A:Reference number: A04627; MUID:77225245; PMID:1142009  
 A:Accession: A04627

A:Molecule type: protein  
 A:Residues: 581-810 <MIL>  
 R:Wiman, B.; Wallen, P.  
 Eur. J. Biochem. 50, 489-494, 1975  
 A:Title: Structural relationship between "glutamic acid" and "lysine" forms of human plas

A:Reference number: A04625; MUID:75093329; PMID:122932  
 A:Accession: A04625  
 A:Molecule type: protein  
 A:Residues: 20-50, 'Q', 51-71, 'E', 73-85, 87-100 <M12>  
 R:Wiman, B.; Wallen, P.  
 Eur. J. Biochem. 58, 539-547, 1975  
 A:Title: Amino-acid sequence of the cyanogen-bromide fragment from human plasminogen that  
 A:Reference number: A04626; MUID:76043692; PMID:126863  
 A:Accession: A04626  
 A:Molecule type: protein  
 A:Residues: 483-507, 'E', 509-604 <M13>  
 R:Robbins, K.C.; Bernabe, P.; Arzadon, L.; Summaria, L.  
 J. Biol. Chem. 248, 1631-1633, 1973  
 A:Title: The primary structure of human plasminogen. II. The histidine loop of human plas  
 A:Reference number: A92125; MUID:73149248; PMID:4694729  
 A:Contents: annotation; active site  
 R:Groekopf, W.R.; Summaria, L.; Robbins, K.C.  
 J. Biol. Chem. 244, 3590-3597, 1969  
 A:Title: Studies on the active center of human plasmin. Partial amino acid sequence of a  
 A:Reference number: A92048; MUID:69234739; PMID:4240117  
 A:Contents: annotation; active site  
 R:Trexler, M.; Vail, Z.; Patchy, L.  
 J. Biol. Chem. 257, 7401-7406, 1982  
 A:Title: Structure of the omega-aminocarboxylic acid-binding sites of human plasminogen.  
 A:Reference number: A92382; MUID:82213905; PMID:6918539  
 A:Contents: annotation; omega-aminocarboxylic acid binding sites  
 J:Vail, Z.; Patchy, L.  
 J. Biol. Chem. 259, 13690-13694, 1984  
 A:Title: The fibrin-binding site of human plasminogen. Arginines 32 and 34 are essential  
 A:Reference number: A92480; MUID:85054794; PMID:6094526  
 A:Contents: annotation; fibrin binding site; omega-aminocarboxylic acid binding site  
 J:Caio, Y.; Ji, R.W.; Davidson, D.; Schaller, J.; Marti, D.; Sehnadel, S.; McCaice, S.G.;  
 J. Biol. Chem. 271, 29461-29467, 1996  
 A:Title: Kringla domains of human angiotensin. Characterization of the anti-proliferative  
 A:Reference number: A58811; MUID:97067211; PMID:8910613  
 A:Contents: annotation  
 R:Uljen, H.R.; Ugwu, F.; Biri, A.; Colten, D.  
 Biochemistry 37, 4699-4702, 1998  
 A:Title: Generation of an angiotensin-like fragment from plasminogen by stromelysin-1 (M  
 A:Reference number: A58812; MUID:9548733; PMID:9548733  
 A:Contents: annotation  
 R:Tullinsky, A.; Mulichak, A.M.  
 submitted to the Brookhaven Protein Data Bank, July 1991  
 A:Reference number: A51341; PDB:1PK4  
 A:Contents: annotation; X-ray crystallography, 1.9 angstroms, residues 376-454  
 R:Tullinsky, A.; Wu, T.P.  
 submitted to the Brookhaven Protein Data Bank, July 1991  
 A:Reference number: A51488; PDB:2PK4  
 A:Contents: annotation; X-ray crystallography, 2.25 angstroms, residues 375-454  
 R:Wu, T.P.; Tullinsky, A.  
 submitted to the Brookhaven Protein Data Bank, August 1993  
 A:Reference number: A51911; PDB:1PKR  
 A:Contents: annotation; X-ray crystallography, 2.48 angstroms, residues 102-181  
 R:Padmanabhan, K.; Tullinsky, A.  
 submitted to the Brookhaven Protein Data Bank, April 1994  
 A:Reference number: A52408; PDB:1PMK  
 A:Contents: annotation; X-ray crystallography, 2.25 angstroms, residues 377-454  
 R:Tullinsky, A.; Mathews, I.I.  
 submitted to the Brookhaven Protein Data Bank, December 1995  
 A:Reference number: A65244; PDB:1CEA  
 A:Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 102-181  
 R:Tullinsky, A.; Mathews, I.I.  
 submitted to the Brookhaven Protein Data Bank, December 1995  
 A:Reference number: A65245; PDB:1CEB  
 A:Contents: annotation; X-ray crystallography, 2.1 angstroms, residues 102-181  
 R:Mulichak, A.M.; Tullinsky, A.; Ravichandran, K.G.  
 Biochemistry 30, 10576-10586, 1991

A/Title: Crystal and molecular structure of human plasminogen kringle 4 refined at 1.9 Å  
A/Reference number: A58819; MUID:92031502; PMID:1657148  
A/Contents: annotation  
R/MU: T.P.; Padmanabhan, K.; Tulinsky, A.; Mulichak, A.M.  
A/Title: The refined structure of the epsilon-aminocaproic acid complex of human plasmin  
A/Reference number: A58818; MUID:92031503; PMID:1657149  
A/Contents: annotation  
R/De Vos, A.M.; Ultsch, M.H.; Kelley, R.F.; Padmanabhan, K.; Tulinsky, A.; Westbrook, M.  
Biochemistry 31, 270-279, 1992  
A/Title: Crystal structure of the kringle 2 domain of tissue plasminogen activator at 2.0 Å  
A/Reference number: A39483; MUID:92118803; PMID:1310033  
A/Contents: annotation; X-ray crystallography, 2.4 angstroms  
R/Steck, B.; Teeter, M.M.; Whitlow, M.; Yamano, A.  
Submitted to the Brookhaven Protein Data Bank, June 1995  
A/Reference number: A65980; PDB:1KRN  
A/Contents: annotation; X-ray crystallography, 1.67 angstroms, residues 376-454  
R/Rejzante, M.; Llinas, M.  
Submitted to the Brookhaven Protein Data Bank, August 1996  
A/Reference number: A65803; PDB:1HPJ  
A/Contents: annotation; conformation by (1)H-NMR, residues 103-181  
R/Rejzante, M.; Llinas, M.  
Submitted to the Brookhaven Protein Data Bank, August 1996  
A/Reference number: A65804; PDB:1HPK  
A/Contents: annotation; conformation by (1)H-NMR, residues 103-181  
R/Rejzante, M.R.; Llinas, M.  
Eur. J. Biochem. 221, 927-937, 1994  
A/Title: (1)H-NMR assignments and secondary structure of human plasminogen kringle 1.  
A/Reference number: S43645; MUID:94237157; PMID:8181475  
A/Contents: annotation; conformation by (1)H-NMR, residues 96-184  
R/Rejzante, M.R.; Llinas, M.  
Eur. J. Biochem. 221, 939-949, 1994  
A/Title: Solution structure of the epsilon-aminohexanoic acid complex of human plasminogen  
A/Reference number: A58817; MUID:94237158; PMID:8181476  
A/Contents: annotation; conformation by (1)H-NMR  
C/Comment: Plasminogen is synthesized by the kidney and is present in plasma and many of  
C/Comment: Plasminogen is converted to plasmin by plasminogen activators (see PIR:UKHU  
d PIR:FGHUGB).  
C/Comment: Plasmin is inactivated by alpha-2-antiplasmin (see PIR:ITHUA2) immediately at  
rg-380, resulting in two chains connected by two disulfide bonds. Without the inhibitor,  
C/Comment: Microplasmin is formed by autocatalytic cleavage of plasmin under artificial cond  
C/Comment: Streptomyisin 1 (see PIR:KCHUS1) acts on plasminogen to produce angiotensin. Tc  
ting solid tumors.  
C/Genetics:  
A/Gene: GDB:PLG  
A/Cross-references: GDB:119498; OMIM:173350  
A/Map position: 6q26-q27  
A/Intons: 17/1; 62/2; 98/1; 136/2; 183/1; 223/2; 263/1; 317/2; 366/1; 419/2; 480/1; 529  
C/Function:  
A/Description: dissolves the fibrin of blood clots; acts as a proteolytic factor in a va  
ne the walls of the graafian follicle; also activates the urokinase-type plasminogen act  
A/Pathway: fibrinolysis  
C/Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homology  
C/Keywords: angiogenesis inhibitor; blood; duplication; fibrinolysis; glycoprotein; hyd  
F:1-96/Domain: plasminogen-related protein precursor homology <PLPH>  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-810/Product: plasminogen #status experimental <PRO>  
F:20-96/Domain: activation peptide #status experimental <APT>  
F:79-466/Product: angiotensin #status experimental <AST>  
F:97-580,581-810/Product: plasmin #status experimental <MAT>  
F:97-580/Domain: plasmin chain A #status experimental <CHA>  
F:103-181/Domain: kringle homology <KR1>  
F:185-262/Domain: kringle homology <KR2>  
F:275-352/Domain: kringle homology <KR3>  
F:377-454/Domain: kringle homology <KR4>  
F:481-560/Domain: kringle homology <KR5>  
F:550-580,581-810/Product: microplasmin #status experimental <MMT>

Query Match 7.5%; Score 101; DB 1; Length 810;  
Best Local Similarity 18.0%; Pred. No. 0.4;  
Matches 53; Conservative 42; Mismatches 90; Indels 110; Gaps 13;  
QY 5 LVGFARVVSADSVHRCISACLNAPDTFGFECESVMYYPVDAECILNTEDR-----LDRP 59

Db 33 LFSVTKKQLGAGSIEECAKCEDEE---FTCRFAFOYHSKEQCQVIMAEKSSIIIRMR 89  
QY 60 DLPLFDEHEDVYIYIDNNCAGCECHWFDPNFTSGILNDQCPALAAQCAPVYTOYVAVG 119  
Db 90 DVVLFPEKK--VYLSECKTGKKNVYRGTMSKTKNGITCQKWSST---SPH----- 133  
QY 120 RQLSDELDHSPFEGLE-----LSECEECCTQ----- 144  
Db 134 RPRSPATHPSBEGLEENYCRNPDPDPOGPMCTTDPDEKRYDCILBCEBECMCISGENY 193  
QY 145 --RLSVTANDFNCKS-----FMYSNLTRS--CVLSDESRPLGRANLAEVP 186  
Db 194 DGIKSKTMGSLGECQAMDSQSPHAGYIPSKRPNNLKKVYCRNPDRDLRP----- 243  
QY 187 GWTFESRGVDFPRVQMLIVGFASFMENVPVMTCLDOCTSPPTPTGQNPVC 241  
Db 244 -W-----CFYTDPN-----KRWELCDIP-----RCTTPPSSSGPTYC 275

RESULT 15  
T25288  
hypothetical protein T26C5.2 - Caenorhabditis elegans  
C/Species: Caenorhabditis elegans  
C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C/Accession: T25288  
R/Thomas, K.  
Submitted to the EMBL Data Library, August 1995  
A/Reference number: Z20011  
A/Accession: T25288  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-554 <MIL>  
A/Cross-references: UNIPROT:Q22815; EMBL:Z50859; PIDD:CAA90727.1; GSPDB:GN00020; CESP:T2K  
A/Experimental source: clone T26C5  
C/Genetics:  
A/Gene: CESP:T26C5.2  
A/Map position: 2  
A/Intons: 20/1; 59/3; 103/1; 266/1; 333/1; 376/3; 542/2  
C/Superfamily: Caenorhabditis elegans hypothetical protein T26C5.2

Query Match 7.3%; Score 99.5; DB 2; Length 554;  
Best Local Similarity 28.8%; Pred. No. 0.35;  
Matches 19; Conservative 11; Mismatches 33; Indels 3; Gaps 1;  
QY 16 DSVHRCISACLNAPDTFGFECESVMYYPVDAECILNTEDRDLPLFVDEHEDVYIYIDN 75  
Db 456 DGMQLCTIELCVLSTK---FTCRSSSTFNPITGQCRLMTEDSWTSPDSFEYDFQKALYPEN 512  
QY 76 NCAGE 81  
Db 513 GCTNAE 518

Search completed: March 31, 2005, 02:06:55  
Job time : 45 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2005, 01:56:09 ; Search time 177 Seconds  
(without alignments)  
723.276 Million cell updates/sec

Title: US-10-771-708-11  
Perfect score: 1354  
Sequence: 1 EQKLIVGFARVVSADSVHR.....PPETGQNPVCKSVMYYYNE 250

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_03:.\*  
1: uniprot\_sprot:.\*  
2: uniprot\_tr embl:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	912.5	67.4	1065	2	001494 caenorhabdi
2	912.5	67.4	1069	2	0833b7 caenorhabdi
3	262	19.4	751	2	020167 caenorhabdi
4	168	12.4	741	2	09XUES caenorhabdi
5	160	11.8	522	2	017347 caenorhabdi
6	157	11.6	1494	2	07PTE3 anopheles g
7	153.5	11.3	398	2	022902 caenorhabdi
8	138	10.2	1549	2	08MKZ3 drosophila
9	138	10.2	1549	2	09BMD4 drosophila
10	138	10.2	1557	2	09BMD5 drosophila
11	138	10.2	1557	2	09V5X3 drosophila
12	126	9.3	351	2	021782 caenorhabdi
13	125	9.2	774	2	09V9X1 drosophila
14	124.5	9.2	1127	2	094248 caenorhabdi
15	121	8.9	744	2	09VAG2 drosophila
16	120.5	8.9	432	2	07PUL3 drosophila
17	120.5	8.9	601	2	09W520 drosophila
18	120.5	8.9	715	2	08MS37 drosophila
19	119.5	8.8	633	2	07QBU4 anopheles g
20	118.5	8.8	695	2	07QDP6 anopheles g
21	116	8.6	790	1	08K667 sus scrofa
22	115	8.5	805	2	081GS2 drosophila
23	113	8.3	833	3	09V9X0 drosophila
24	108	8.0	812	1	09V9X0 drosophila
25	105	7.8	812	1	PLMN_BOVIN
26	104	7.7	810	1	PLMN_MACMU
27	103	7.6	498	1	YSMS_CABEU
28	102.5	7.6	942	1	Q18298 caenorhabdi
29	101	7.5	810	1	PLMN_HUMAN
30	99.5	7.3	524	2	Q22815 caenorhabdi
31	99	7.3	466	2	Q6RC10 mus musculu

32	99	7.3	510	2	060123 echizosacch
33	99	7.3	693	2	Q27394 caenorhabdi
34	99	7.3	812	1	PLMN_MOUSE
35	98	7.2	1580	2	07RME4 plasmodium
36	96.5	7.1	806	1	PLMN_MACMU
37	96	7.1	627	2	064713 arbidopais
38	96	7.1	627	2	064716 arbidopais
39	96	7.1	627	2	084LS0 arbidopais
40	96	7.1	824	2	066S04 oikopleura
41	95.5	7.1	696	2	061814 caenorhabdi
42	94.5	7.0	1615	2	07QZU9 giardia lam
43	94	6.9	429	2	072206 bacillus ce
44	94	6.9	786	2	062201 caenorhabdi
45	93.5	6.9	480	2	Q6GZL3 bartonella

ALIGNMENTS

RESULT 1  
001494 PRELIMINARY; PRT; 1065 AA.  
ID 001494  
AC 001494;  
DT 01-JUL-1997 (TREMBLrel. 04, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Hypothetical protein C34G6.6;  
GN Name=C34G6.6; ORFNames=C34G6.6;  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Pelodierinae; Caenorhabditis.  
NC NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RX MEDLINE=99069613; PubMed=9851916;  
RG WormBase Consortium;  
RT "Genome sequence of the nematode C. elegans: a platform for  
investigating biology. The C. elegans Sequencing Consortium.,"  
RL Science 282:2012-2018 (1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Langston Y., Rohlfing T.;  
RT "The sequence of C. elegans cosmid C34G6.6";  
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RA Waterston R.;  
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RG WormBase Consortium;  
RL Submitted (SRP-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U97407; AAB52479.2; -.  
DR PIR; C87789; C87789.  
DR WormBase; WBGene0016422; C34G6.6.  
DR WormPep; C34G6.6a; CE29699.  
DR InterPro; IPR001507; Endoglin/CD105.  
DR InterPro; IPR003014; PAN.  
DR InterPro; IPR003609; Pan\_app.  
DR Pfam; PF00024; PAN; 6.  
DR Pfam; PF00100; Zona\_pellucida; 1.  
DR SMART; SMO0473; PAN\_AP; 6.  
DR SMART; SMO0241; ZP; 1.  
DR PROSITE; PS50948; PAN; 6.  
KW Hypothetical protein.  
SQ SEQUENCE 1065 AA; 117427 MW; DB8597B650648F84 CRC64;  
Query Match 67.4%; Score 912.5; DB 2; Length 1065;  
Best Local Similarity 66.2%; Pred. No. 8.6e-76;

	Matches	172:	Conservative	33;	Mismatches	36;	Indels	19;	Gaps	5					
Qy	1	EOKLIVGAREVVSADSVHRCISACLNAPDTFGFECESVMYYPVDAECLINTEDLDPRD	60												
Dy	114	EORLLVGFAREVVPAAANIQQICMAACLINAFDTFGFECESAMFPVDOECILNTEDLDPRS	173												
Qy	61	LFEVDEHEETTVYLDDNCCAGCSCHMHFDPAFKTSGILINDQOFAIA--AOCAPVTOTYVAEG	119												
Dy	174	LFEVESDELDTVIYMDNNCAGFLP--VFKNY-----NYKTTFSKSQCYPPIYTOTYIAVEG	225												
Qy	120	ROLSDELDPHSSEGLELSECBEICTQRLSVTANDFNCKSFMYSNLTRSCVLSDERSRPIGR	179												
Dy	226	KQLKNELDRIT-NVDLDSGCQALCTQRSSISSNDFNCKSFMNNKTKRTGLIADERSKPLGR	284												
Qy	180	ANLAEVGPWTFYES-----RGVPSFTFRVPMQLLVGFASPVMEAVPSVTMLCDQCTS	230												
Dy	285	ADLTATSGFTTFEFKKCFSPSPMTCHENVPSFKFVPQMLLVGFAPVMEAVPSVTMLCDQCTN	344												
Qy	231	PPPETGONFVCKSVMYYYNE	250												
Dy	345	PPPETGDGFVCKSVMYYYNE	364												
<hr/>															
	RESULT 2														
ID	Q8T3B7	PRELIMINARY;	PRT,	1069	AA.										
AC	Q8T3B7;														
DT	01-JUN-2002 (TREMBLrel. 21, Created)														
DT	01-JUN-2002 (TREMBLrel. 21, Last sequence update)														
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)														
DE	Hypothetical protein C34G6.6;														
GN	Name=C34G6.6; ORFNames=C34G6.6;														
OS	Caenorhabditis elegans.														
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditiida; Rhabditoidea;														
CC	Rhabditiidae; Peloderinae; Caenorhabditis.														
OX	NCBI_TaxID=6239;														
RN	[1]														
RP	SEQUENCE FROM N.A.														
RC	STRAIN=Bristol N2;														
RG	MEDLINE=99069613; PubMed=9851916;														
RT	WormBase Consortium;														
RL	"Genome sequence of the nematode C. elegans: a platform for investigating biology. The C. elegans Sequencing Consortium."; Science 282:2012-2018 (1998).														
RN	[2]														
RP	SEQUENCE FROM N.A.														
RC	STRAIN=Bristol N2;														
RA	Langston Y., Rohlfing T.;														
RL	"The sequence of C. elegans cosmid C34G6.";														
RN	Submitted (Apr-1997) to the EMBL/GenBank/DBJ databases.														
RP	[3]														
RC	SEQUENCE FROM N.A.														
RA	STRAIN=Bristol N2;														
RL	Waterston R.;														
RN	Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.														

Query Match	Similarity	67.4%	Score 912.5	DB 2	Length 1069
Best Local Similarity	66.2%	Pred. No. 8.6e-76			
Matches 172	Conservative 33	Mismatches 36	Indels 19	Gaps 5	
Qy	1	EQKILVGFARVVSADSVHRCLSAACLNAPDTPFGECESVMYYPVDAECILNTEDRLDRPD	60		
Db	114	EQKILVGFARVVSADSVHRCLSAACLNAPDTPFGECESVMYYPVDAECILNTEDRLDRPD	173		
Qy	61	LFVDEHEBTVIYLLNNCCGCGCHHFPDNFKTSGILNDQFALA-AQCAPVYQYVAVEG	119		
Db	174	LFVSESDDTVLYMNNCCGFPPL-VFNKY-----NFKTFSKSCYCPPIYQYIAVEG	225		
Qy	120	QSLSELDHSEFGEGLSECEELCTQRLSVTANDFNCKSPMYSNLTSCVLSDERSPLGR	179		
Db	226	QKLNKNEIDRLIT-NVDDLSGCMALCTQRSLISSNDPNCKSPMYNNKTRICILADESKPLGR	284		
Qy	180	ANLAEVDPGWYFES-----RGVPSFTRVPOMLLVGFASPMENVPSTVMCLDQCTS	230		
Db	285	ADLIATSGFTYFEKKCFASPTTCGNVPSFKRVPQMLVGFAPMENVPSTVMCLDQCTN	344		
Qy	231	PPPETGQNFVCKSVMYTYNE 250			
Db	345	PPPETGDFVCKSVMYTYNE 364			
RESULT 3					
ID	Q20167	PRELIMINARY	PRT	751 AA.	
AC	Q20167				
DT	01-NOV-1996	(TREMBlrel. 01, Created)			
DT	01-NOV-1996	(TREMBlrel. 01, Last sequence update)			
DT	01-JUN-2003	(TREMBlrel. 24, Last annotation update)			
DE		Hypothetical protein F38E11.4.			
GN		ORFNames=F38E11.4;			
OS		Caenorhabditis elegans.			
OC		Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;			
CC		Rhabditidae; Peloderinae; Caenorhabditis.			
CC		NCBI_TaxID=6239;			
RN	[1]				
RP		SEQUENCE FROM N.A.			
RC		STRAIN=Bristol N2.			
RX		MEDLINE=99069613; PubMed=9851916;			
RA		none;			
RT		"Genome sequence of the nematode C.elegans: A platform for			
RL		investigating biology.";			
RL		Science 282:2012-2018(1998).			
RN	[2]				
RP		SEQUENCE FROM N.A.			
RC		STRAIN=Bristol N2;			
RA		Matthews P.;			
RL		Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.			
RL		EMBL; Z68342; CAA92773.1; -.			
DR	PIR; T21967;				
DR	WormBase; WBGene00009541; F38E11.4.				
DR	WormPeP; F38E11.4; C003296.				
DR	InterPro; IPR001507; Endoglin/CD105.				
DR	InterPro; IPR003014; PAN.				
DR	Pfam; PF00024; Pan_4.				
DR	Pfam; PF00100; Zona pellucida; 1.				
DR	SMART; SM00473; PAN_Ap; 4.				
DR	SMART; SM00241; ZP_1.				
DR	PROSITE; PSS0948; PAN; 4.				
DR	Hypothetical protein.				
QY	SEQUENCE 751 AA; 92AEFE353CCCBFA4 CRC64;				
Query Match	19.4%	Score 262;	DB 2;	Length 751;	
Best Local Similarity	27.4%	Pred. No. 1.7e-15;			
Matches 73	Conservative 37	Mismatches 98	Indels 58	Gaps 8	
QY	2	QKILVGFARVVSADSVHRCLSAACLNAPDTPFGECESVMYYPVDAECILNTEDRLDRPD	61		

```

Db      131 OSILIGHAMKVLTVYDGLSDCLSKALSORSYDFLCKSAIYYETGECIMRDNKFIYPKL 190
Qy      62 FVDEHEDVLT-YLDNNCGGEC-----HHFPDNFKTSGILNDOQFAIAOQVAPYVQY 114
Db      191 FKTILDLTVDFENNCADVSCKEPETHM-----220
Qy      115 VAVGRQSLDLDHSEFGLSECEBELCTQRLSTVANDFNCKSPMYSNLTRS CVLSDRS 174
Db      221 VRTEBYLIDSKDVIYESSDAOECNQLC-QNNKIGENFPCKAPAYSNKQECHLITAESS 279
Qy      175 ----RPLGRANLAEPVGMWTFESRGVP-----SFTRVPMQLLVGFASFVMEVPSVT 222
Db      280 YVGHKGDRFNLAPLNPSEGEYFEKYLPTNLQICIASFELVANRMMWT--SAVKTISALSQH 337
Qy      223 MCLDOCTSPPEPTGQNFVCKSVMYTY 248
Db      338 ECLSQCMNDGAR-----CSSATYFY 357

RESULT 4
Q9XUE5 PRELIMINARY; PRT; 741 AA.
ID      09XUE5
AC      09XUE5;
DT      01-NOV-1999 (TREMBLrel. 12, Created)
DT      01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT      01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE      Hypothetical protein F52B11.3.
GN      ORFNames=F52B11.3;
OS      Caenorhabditis elegans.
OC      Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
OC      Rhabditidae; Pelodertinae; Caenorhabditis.
OX      NCBI_TaxId=6239;
RN      [1]
RC      SEQUENCE FROM N.A.
RC      STRAIN=Bristol N2;
RX      MEDLINE=99069613; PubMed=9851916;
RA      none;
RT      "Genome sequence of the nematode C.elegans: A platform for
RT      investigating biology.";
RL      Science 282:2012-2018(1998).
RN      [2]
RC      SEQUENCE FROM N.A.
RC      STRAIN=Bristol N2;
RA      Matthews L.;
RA      Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RL      EMBL; Z82486; CAB05199.1; -.
DR      PIR; T22486; T22486.
DR      WormBase; MBGene0000926; F52B11.3.
DR      WormPeP; F52B11.3; CE18723.
DR      InterPro; IPR001507; Endoglin/CD105.
DR      InterPro; IPR003014; PAN.
DR      Pfam; PF00024; PAN; 4.
DR      Pfam; PF00100; zona pellucida; 1.
DR      SMART; SM00473; PAN_AP; 4.
DR      SMART; SM00241; ZP; 1.
DR      PROSITE; PS50948; PAN; 4.
DR      Hypothetical protein.
SQ      SEQUENCE 741 AA; 81738 MW; F6BC96FED7D8C15 CRC64;

Query Match 12.4%; Score 168; DB 2; Length 741;
Beet Local Similarity 22.5%; Pred. No. 9.3e-07;
Matches 76; Conservative 38; Mismatches 100; Indels 124; Gaps 15;

5 LVGFARVVSADSVARCSACIAPDITGFECESVMYVVDACIANTEDRLDRDLPVD 64
137 LEGFVKSVTVENRHCISACKEK-----FVCKSVNPHDTSICELSVEDKSKRP----- 188
65 EH---EDVTYIYDNNCAG-----C--EC 82
189 THVRMSKIDYDNNCLSRONCGPSGNLVVKTTFEIRYDHTQSVAEQSYCLQKC 248
83 HHFPDNFKTSGILN-----DQ-----FALIAQCYAPYVQY 114

```

```

Db      249 LDSLNTFCRSVFEVPEKNCIVSDEDTFSRADQGOVGVKDYEPICVAADLSSTCRQ 308
Qy      115 VAVE---GRQLSDLDHSEFGLSECEBELCTQRLSTVANDFNCKSPMYSNLTRS CVLSD 171
Db      309 AAFERFIGSSIEGVVVAQAQVTTISDCLCFQNL-----NCKSINVDRTASSCFI-- 359
Qy      172 ERSRPLGR--ANLAEPVGMWTFESRGVPSF-----TRVPMQLLVGFA-- 211
Db      360 ---YAVGRQDNIYANDSMYIEFNCSQFGMALCTNEGIRFTVNTKEPYTGAIYAAR 416
Qy      212 ----SFVMEVPSVTMCLDOCTSPPEPTGQNFVCKSVY 245
Db      417 FSTCSQVVENAKQISL-----TFPPPTVSSD--CGTVI 447

RESULT 5
O17347 PRELIMINARY; PRT; 522 AA.
ID      017347
AC      017347;
DT      01-JAN-1998 (TREMBLrel. 05, Created)
DT      01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT      01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE      Hypothetical protein H42K12.3.
GN      Name=H42K12.3; ORFNames=H42K12.3;
OS      Caenorhabditis elegans.
OC      Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
OC      Rhabditidae; Pelodertinae; Caenorhabditis.
OX      NCBI_TaxId=6239;
RN      [1]
RC      SEQUENCE FROM N.A.
RC      STRAIN=Bristol N2;
RX      MEDLINE=99069613; PubMed=9851916;
RG      WormBase Consortium;
RT      "Genome sequence of the nematode C. elegans: a platform for
RT      investigating biology. The C. elegans Sequencing Consortium.";
RL      Science 282:2012-2018(1998).
RN      [2]
RC      SEQUENCE FROM N.A.
RC      STRAIN=Bristol N2;
RA      Maggi L., Harper M.;
RA      "The sequence of C. elegans fosmid H42K12.";
RL      Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
RN      [3]
RC      SEQUENCE FROM N.A.
RC      STRAIN=Bristol N2;
RA      Waterston R.;
RL      Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RN      [4]
RC      SEQUENCE FROM N.A.
RC      STRAIN=Bristol N2;
RA      Wilson R.;
RL      Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
RN      [5]
RC      SEQUENCE FROM N.A.
RC      STRAIN=Bristol N2;
RG      WormBase Consortium;
RL      Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
RL      EMBL; AF026207; AAB71266.1; -.
DR      PIR; T32444; T32444.
DR      WormBase; MBGene00019272; H42K12.3.
DR      WormPeP; H42K12.3; CE11614.
DR      InterPro; IPR003014; PAN.
DR      InterPro; IPR003609; Pan_APP.
DR      Pfam; PF00024; PAN; 2.
DR      Pfam; PF00084; TSP1.
DR      SMART; SM00473; PAN_AP; 2.
DR      SMART; SM00209; TSP1; 1.
DR      PROSITE; PS50948; PAN; 2.
DR      PROSITE; PS50092; TSP1; 1.
DR      Hypothetical protein.
SQ      SEQUENCE 522 AA; 58738 MW; C4E804F9A0059180 CRC64;

```



DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 DE CG13207-PC (CG13207-PG).  
 GN Name=nompA; ORFNames=CG13207;  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Peerygota;  
 CC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NCBI\_Taxid=7227;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Venter L., Zhang Q., Chent L.X.,  
 RA Brandon R.C., Rogers Y.H., Blazej R.G., Champagne M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gaber G.L.,  
 RA Abail J.F., Agbayani A., An H.J., Andrews-Pfankoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Bertman B.P., Bhandari D., Bolshakov S.,  
 RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Broctier P.,  
 RA Butris K.C., Buesam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Daventport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriere S., Fleischmann W.,  
 RA Foeller C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaeser K.,  
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,  
 RA Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Kerchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulj D., Lai Z.,  
 RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mallet B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Mekulov G., Milhina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D., Schaefer F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svrtkac R., Tecor C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.Y., Weissman D.A., Weissstock G.M., Weissbach J.,  
 RA Williams S.M., Woodgett, Morley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
 RA Yen R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Gibbs R.A., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA "The genome sequence of Drosophila melanogaster.";  
 RT Science 287:2185-2195 (2000).  
 RL [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426065; PubMed=12537568;  
 RA Celniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,  
 RA Patel S., Adams M., Champagne M., Dugan S.P., Fries E., Hodson A.,  
 RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,  
 RA Pacle J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
 RA Svrtkac R., Taber P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
 RA Weissstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.,  
 RA "Finishing a whole-genome shotgun Release 3 of the Drosophila  
 RT melanogaster euchromatic genome sequence.";  
 RT Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
 RL [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426070; PubMed=12537573;  
 RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J., Svrtkac R.,  
 RA Patel S., Fries E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
 RA Ashburner M., Celniker S.E.,  
 RA "The transposable elements of the Drosophila melanogaster euchromatin:  
 RT a genomic perspective.";  
 RT Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).  
 RL [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426069; PubMed=12537572;

RA Miara S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bertman B.P.,  
 RA Betencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter C., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.,  
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a  
 RT systematic review.";  
 RT Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 RL [5]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Venter L., Zhang Q., Chent L.X.,  
 RA Brandon R.C., Rogers Y.H., Blazej R.G., Champagne M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gaber G.L.,  
 RA Abail J.F., Agbayani A., An H.J., Andrews-Pfankoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Bertman B.P., Bhandari D., Bolshakov S.,  
 RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Broctier P.,  
 RA Butris K.C., Buesam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Daventport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriere S., Fleischmann W.,  
 RA Foeller C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaeser K.,  
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,  
 RA Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Kerchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulj D., Lai Z.,  
 RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mallet B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Mekulov G., Milhina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D., Schaefer F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svrtkac R., Tecor C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.Y., Weissman D.A., Weissstock G.M., Weissbach J.,  
 RA Williams S.M., Woodgett, Morley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
 RA Yen R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Gibbs R.A., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA "The genome sequence of Drosophila melanogaster.";  
 RT Science 287:2185-2195 (2000).  
 RL [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426065; PubMed=12537568;  
 RA Celniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,  
 RA Patel S., Adams M., Champagne M., Dugan S.P., Fries E., Hodson A.,  
 RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,  
 RA Pacle J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
 RA Svrtkac R., Taber P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
 RA Weissstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.,  
 RA "Finishing a whole-genome shotgun Release 3 of the Drosophila  
 RT melanogaster euchromatic genome sequence.";  
 RT Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
 RL [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426070; PubMed=12537573;  
 RA Kaminker J.S., Bergman C.M., Krommiller B., Carlson J., Svrtkac R.,  
 RA Patel S., Fries E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
 RA Ashburner M., Celniker S.E.,  
 RA "The transposable elements of the Drosophila melanogaster euchromatin:  
 RT a genomic perspective.";  
 RT Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).  
 RL [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426069; PubMed=12537572;

Query Match 10.2%; Score 138; DB 2; Length 1549;  
 Best Local Similarity 21.2%; Pred. No. 0.0014;  
 Matches 58; Conservative 43; Mismatches 100; Indels 72; Gaps 13;  
 3 KIIVGFARREVSASVHRKLSACANAFDTFGFECESVWYVVDACINTEDRLDPLF 62  
 164 KLPISDIKEITAAAR-SDCEKCLN--EFSFVCRSANFSDTMSCTLSRTTRNRPLM 219  
 63 VDEHEDTVYLDNNCAGCECHMFDFNFTSGILNDQOFAIAQCYAPVTVYVAVGROL 122  
 220 EDDNSD-YLENTCLNNE-----RCGCG-LAVFVKENRKL 253  
 123 SD-ELDHSFGLLSBCEBELCTRLSTANDFNCKSPMYSNLRSCLVLSBRS-RPLGR 179  
 254 GGPFEVD-IFNNMTLEECOTWC-----LRAEKYCRSVFEPDQSKCILSEDSISQKOD 307  
 180 ANLAEVPGWTFE-----SRGVSPSTRVPOMLVG-FASF 213  
 308 ISSSSPTHHFHYDLVCLDNOGRANDYPDNTSHLPSSGRRPDTAFORVRSRLGGEFHS 367  
 214 VMENVSVTWCLDOCTSPPEPTGONFVCKSVMY 246  
 368 ITGR--SLSECLDECLRG-----TSFGCRSAVY 393

RESULT 9  
 Q9BMD4 PRELIMINARY; PRT; 1549 AA.  
 AC Q9BMD4;  
 DT 01-JUN-2001 (TREMBlrel. 17, Created)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 DE No-mechanoreceptor potential A short isoform precursor.  
 GN Name=nompA;  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Peerygota;  
 CC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NCBI\_Taxid=7227;  
 [1]

```

RP SEQUENCE FROM N.A.
RX MEDLINE=21134724; PubMed=11239432; DOI=10.1016/S0896-6273(01)00215-X;
RA Chung Y.D., Zhu J., Han Y., Kernan M.J.;
RT "nompA encodes a PMS-specific, ZP domain protein required to connect
RT mechanosensory dendrites to sensory structures.";
RL Neuron 23:415-428(2001).
DR EMBL: AF34032; AAK09434.1; -.
DR FlyBase: FBgn0016047; nompA.
DR GO: GO:0016358; P:dendrite morphogenesis; TAS.
DR GO: GO:0009592; P:detection of sound; IMP.
DR GO: GO:0007605; P:perception of sound; IMP.
DR InterPro: IPR001507; Endoglin/CD105.
DR InterPro: IPR003014; PAN.
DR InterPro: IPR003609; Pan_app.
DR Pfam: PF00024; PAN; 4.
DR Pfam: PF00100; Zona pellucida; 1.
DR SMART: SM00473; PAN_AP; 4.
DR SMART: SM00241; ZP; 1.
DR PROSITE: PS50948; PAN; 5.
DR Receptor; Signal.
KW SIGNAL.
FT SIGNAL.
SQ SEQUENCE 1549 AA; 174270 MW; 78C31BACA39D5B93 CRC64;

Query Match 10.2%; Score 138; DB 2; Length 1549;
Best Local Similarity 21.2%; Pred. No. 0.0014;
Matches 58; Conservative 43; Mismatches 100; Indels 72; Gaps 13;

QY 3 KILVGFAREVVASDSVHRCISACLNAPDTPFGFECESVMYYPVDAECILNTEDRLDRPDLF 62
DB 164 KLPISIDIKETITANR--SDCEDKCLN---EFSFVCRSANFDSITMRSCITLSRFTRRHPETLM 219
QY 63 VDEHEDTVIYLDNNCAGCECHMHFDPNFKTSGILNDQFAIAACVAPYVTOYVAVEGRQL 122
DB 220 EDDPNSD--YLENTCLNAB-----RRCDG--LAVFKEENKRL 253
QY 123 SD--ELDHSFEGLESECEELCTQRLSYTANDFNCKSFMYNSLTRSCVLSDERS--RPLGR 179
DB 254 GGFPEVD--IFNNMTLEECQTM-----LRAEKYFRSVFDPDQSKQCLISEBDSISQKDD 307
QY 180 ANLAEPVPCWTFPE-----SRGVSPFTRVPQMLLVG--FASF 213
DB 308 ISISSPFTHFYDVLCLNDQANDYPDNSVTSHTLSSGRRPDTAFGRYNSRLGGEFHSF 367
QY 214 VMENVPSVTMCLDOCTSPPEPGONFVCKSVMY 246
DB 368 ITGR--SLSECLDCLRQ-----TSFQCRSAVY 393
QY 368 ITGR--SLSECLDCLRQ-----TSFQCRSAVY 393

RESULT 10
Q9BMD5 PRELIMINARY; PRT; 1557 AA.
AC Q9BMD5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE No-mechanoreceptor potential A long isoform precursor.
GN Name=nompA;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxId=7227;
RN NCB1_TaxId=7227;
RP SEQUENCE FROM N.A.
RX MEDLINE=21134724; PubMed=11239432; DOI=10.1016/S0896-6273(01)00215-X;
RA Chung Y.D., Zhu J., Han Y., Kernan M.J.;
RT "nompA encodes a PMS-specific, ZP domain protein required to connect
RT mechanosensory dendrites to sensory structures.";
RL Neuron 23:415-428(2001).
DR EMBL: AF34032; AAK09434.1; -.
DR FlyBase: FBgn0016047; nompA.
DR GO: GO:0016358; P:dendrite morphogenesis; TAS.
DR GO: GO:0009592; P:detection of sound; IMP.
DR GO: GO:0007605; P:perception of sound; IMP.
DR InterPro: IPR001507; Endoglin/CD105.
DR InterPro: IPR003014; PAN.
DR InterPro: IPR003609; Pan_app.
DR Pfam: PF00024; PAN; 4.
DR Pfam: PF00100; Zona pellucida; 1.
DR SMART: SM00473; PAN_AP; 4.
DR SMART: SM00241; ZP; 1.
DR PROSITE: PS50948; PAN; 5.
DR Receptor; Signal.
KW SIGNAL.
FT SIGNAL.
SQ SEQUENCE 1549 AA; 174270 MW; 78C31BACA39D5B93 CRC64;

```

```

DR GO: GO:0007605; P:perception of sound; IMP.
DR InterPro: IPR001507; Endoglin/CD105.
DR InterPro: IPR003014; PAN.
DR InterPro: IPR003609; Pan_app.
DR Pfam: PF00024; PAN; 4.
DR Pfam: PF00100; Zona pellucida; 1.
DR SMART: SM00473; PAN_AP; 4.
DR SMART: SM00241; ZP; 1.
DR PROSITE: PS50948; PAN; 5.
DR Receptor; Signal.
KW SIGNAL.
FT SIGNAL.
SQ SEQUENCE 1557 AA; 175290 MW; 134BBBA89245183 CRC64;

Query Match 10.2%; Score 138; DB 2; Length 1557;
Best Local Similarity 21.2%; Pred. No. 0.0014;
Matches 58; Conservative 43; Mismatches 100; Indels 72; Gaps 13;

QY 3 KILVGFAREVVASDSVHRCISACLNAPDTPFGFECESVMYYPVDAECILNTEDRLDRPDLF 62
DB 164 KLPISIDIKETITANR--SDCEDKCLN---EFSFVCRSANFDSITMRSCITLSRFTRRHPETLM 219
QY 63 VDEHEDTVIYLDNNCAGCECHMHFDPNFKTSGILNDQFAIAACVAPYVTOYVAVEGRQL 122
DB 220 EDDPNSD--YLENTCLNAB-----RRCDG--LAVFKEENKRL 253
QY 123 SD--ELDHSFEGLESECEELCTQRLSYTANDFNCKSFMYNSLTRSCVLSDERS--RPLGR 179
DB 254 GGFPEVD--IFNNMTLEECQTM-----LRAEKYFRSVFDPDQSKQCLISEBDSISQKDD 307
QY 180 ANLAEPVPCWTFPE-----SRGVSPFTRVPQMLLVG--FASF 213
DB 308 ISISSPFTHFYDVLCLNDQANDYPDNSVTSHTLSSGRRPDTAFGRYNSRLGGEFHSF 367
QY 214 VMENVPSVTMCLDOCTSPPEPGONFVCKSVMY 246
DB 368 ITGR--SLSECLDCLRQ-----TSFQCRSAVY 393
DB 368 ITGR--SLSECLDCLRQ-----TSFQCRSAVY 393

RESULT 11
Q9V5X3 PRELIMINARY; PRT; 1557 AA.
AC Q9V5X3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE CG13207-PA (CG13207-pb).
GN Name=nompA; ORFNames=CG13207;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxId=7227;
RN NCB1_TaxId=7227;
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.D.,
RA George R.A., Lewis S.B., Richards S., Ashburner M., Henderson S.N.,
RA Sutcliffe G.G., Wortman J.R., Yandell M.D., Zhang G., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abail J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodeson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durkin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrieli A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,

```

RA Hostlin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,  
RA Jaitani M., Kalush P., Karpen G.H., Ke Z., Kennison J.A., Kelchum K.A.,  
RA Klimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Liao X., Matei B., McIntosh A.C., Li Y., Li Z., Liang Y., Lin X.,  
RA Mekulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusserm D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Fui V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skusek M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svaydas R., Teclor C., Turner R., Ventler E., Wang A.H., Wang X.,  
RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodagert, Morley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
RA Yen R.F., Zaveri U.S., Zhao M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of *Drosophila melanogaster*";  
RL Science 287:2165-2195(2000).  
[2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22426065; PubMed=12537568;  
RA Centker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,  
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,  
RA George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,  
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.,  
RT "Finishing a whole-genome shotgun: Release 3 of the *Drosophila*  
melanogaster euchromatic genome sequence";  
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
[3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22426070; PubMed=12537573;  
RA Kimkiner J.S., Bergman C.M., Kronmiller B., Carlson J., Svaydas R.,  
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
RA Ashburner M., Centker S.E.;  
RT "The transposable elements of the *Drosophila melanogaster* euchromatin:  
a genomics perspective.";  
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).  
[4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22426069; PubMed=12537572;  
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
RA Hardecky P., Huang Y., Kimmin J.S., Millburn G.H., Prochnik S.E.,  
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglou L., Berman B.P.,  
RA Bettencourt B.R., Centker S.E., de Grey A.D., Drysdale R.A.,  
RA Harris N.L., Richter U., Russo S., Schroeder A.J., Shu S.Q.,  
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
Lewis S.E.;  
RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
systematic review";  
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
[5]  
RP SEQUENCE FROM N.A.  
RG FlyBase;  
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
[6]  
RP SEQUENCE FROM N.A.  
RG FlyBase;  
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
EMBL: AEO03826; AA58668.2; -  
FlyBase: FBgn0016047; nompA.  
DR GO:GO:0016358; P:dendrite morphogenesis; TAS.  
DR GO:GO:0009592; P:detection of sound; IMP.  
DR GO:GO:0007605; P:piperception of sound; IMP.  
DR InterPro: IPR001507; Endoglin/CD105.  
DR InterPro: IPR003609; Pan\_app.  
DR Pfam: PF00024; PAN\_4.  
DR Pfam: PF00024; PAN\_4.  
DR SMART: SM00473; PAN\_AF\_4.

DR	SMART; SMO0241; ZP; 1.
DR	PROSITE; PS50949; PAN; 5.
SO	SEQUENCE 1557 AA; 17526 MW; E060906D48DB77DF CRCK4;
Query Match	10.2%; Score 138; DB 2; Length 1557;
Best Local Similarity	21.2%; Pred. No. 0.0014;
Matches	58; Conservative 43; Mismatches 100; Indels 72; Gaps 13
DQ	3 KILVGPAREVVSADSVHRCLSACLNAPDTFGEGECSSVMYTPVDACEILTNERLDRDPDLF 62
Dd	164 KLPISDIKETIAANR--SPDCBKCLN---EFSEVCRSANAFDSITMSRCTLSRFRRTHPELM 219
Dy	63 VDEHDHYIYLIDNNCAGCECHMHNDPFKTSGLINDQCALMAQCAPVTQVVAATEGROL 122
Dd	220 EDDPNSD--YLENTCLNAE-----RCRG--LAIVFKENKRL 253
Dy	123 SD--ELHDSFEGLISECEELECTORLSTVANDFNCKSFMYNLTRSCVLSDERS-RPLGR 179
Dd	254 GGPEFVD-IENNMTLEBCQTWC---LRACKYFCRSVEFPDQSQCGLTSEBDSISQMD 307
Dy	180 ANLAIEVGWTFE-----SRGVPSRTVPQMLLVG-FASF 213
Dd	308 ISISSPTFHMFYDVLVCLDNCRANDYPDNSVYSHLFSSSGRRPDTAFORYNRSRLGGEPFSH 367
Dy	214 VMENVPSVTMGLDOCTSPPETGONPVFKSVMY 246
Dd	368 ITGR-SLSECLDECLRO----TSFOCRSAVY 393

RESULT 12	021782	PRELIMINARY;	PRT,	351 AA.
ID	Q21782.			
AC	Q21782.			
DT	01-NOV-1996 (TrEMBLrel. 01, Created)			
DT	01-NOV-1996 (TrEMBLrel. 01, Last sequence update)			
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)			
DE	Hypothetical protein R07A4.4.			
GN	ORFNames=R07A4.4;			
OS	Caenorhabditis elegans.			
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabditoidea;			
OC	Rhabdilitidae; Peloiderinae; Caenorhabditis.			
OX	NCBI_TaxId=6239;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Bristol N2;			
KX	MEDLINE=99069613; PubMed=9851916;			
RA	none;			
RT	"Genome sequence of the nematode C.elegans: A platform for			
RT	investigating biology.";			
RL	Science 282:2012-2018(1998).			
RL	[2]			
RN	SEQUENCE FROM N.A.			
RP	STRAIN=Bristol N2;			
RC	Cottage A.;			
RL	Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; Z67756; CAA91764.1; -.			
DR	PIR; T23990; T23990.			
DR	WormBase; WBGene00011075; R07A4.4.			
DR	WormPep; R07A4.4; CE03533.			
DR	InterPro; IPR003014; PAN.			
DR	InterPro; IPR003609; Pan_app.			
DR	InterPro; IPR008266; Tyr_Dkinase_AS.			
DR	Pfam; PF00024; PAN.2.			
DR	SMART; SMO00473; PAN.Ap.2.			
DR	PROSITE; PSS0948; PAN.2.			
DR	PROSITE; PS00109; PROTEIN_KINASE_TYR; UNKNOWN_1.			
KW	Hypothetical protein.			
SQ	SEQUENCE 351 AA; 39349 MW; C9E0EBCC3612FB0 CRC64;			
Query March	9.3%;	Score 126;	DB 2;	Length 351;
Best Local Similarity	25.0%;	Pred. No. 0.0029;		
Matches 45;	Conservative 31;	Mismatches 72;	Indels 32;	Gaps 7

QY 95 LINDQ-----FA-----IAQCYAPVYQVAVAGEHQLSDHDSFEGLESECEBL 141  
 DB 8 INQOQOQFNDTALNFPAVEVLPIVDDCPLGHSRQVITEGVESEALITITQVAILLEQVIA 67  
 QY 142 CTQRLSVLTANDFN-----CKSFMYSNLTRSCVLSDESRRLGRANLAIEVGMWTFESRGVP 197  
 DB 68 C-RVSTVADSGSRLLPLGCSAHPNRATRCQSVSDAINFNGLVLEYKPNQNIYIEKICIP 125  
 QY 198 S-----FTRVQMLLVGFASFVMEVNSVTMCLDOQCSPPRETOGNFVCSVMY 247  
 DB 126 DTVLPWSCDVPRRIPQHILGLHASEVI-SVASEMECVLECIX-AKTLRSVACHSILHY 182  
 RESULT 13  
 QY9X1 PRELIMINARY; PRT; 774 AA.  
 ID QY9X1  
 AC QY9X1;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE CG12063-PA.  
 GN ORFNames=CG12063;  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NCBI\_TaxID=7227;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amaratides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe W., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gaber G.L.,  
 RA Abiril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballwey R.M., Basu N.A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolhaver S.,  
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,  
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasner K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,  
 RA Jajalil M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Mlishina N.V., Mobarry C., Morris J., Mostrel A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D., Scheel J., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,  
 RA Williams S.M., Woodaght, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426065; PubMed=12537568;  
 RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,  
 RA Patel S., Adams M., Champe W., Dugan S.P., Frise E., Hodgson A.,

RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,  
 RA Pacle J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
 RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
 RA Weisskopf G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.,  
 RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila  
 RT melanogaster euchromatic genome sequence.";  
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426070; PubMed=12537573;  
 RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,  
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
 RA Ashburner M., Celniker S.E.;  
 RT "The transposable elements of the Drosophila melanogaster euchromatin:  
 RT a genomic perspective.";  
 RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Milburn G.H., Prochuk S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
 RA Belencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Ruso S., Schroeder A.J., Shu S.Q.,  
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RG Flybase;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RG Flybase;  
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AE003778; AAF5158.3;  
 DR Flybase; FBgn0039851; CG12063.  
 DR InterPro; IPR001507; Endoglin/CD105.  
 DR InterPro; IPR003014; PAN.  
 DR InterPro; IPR003609; Pan\_app.  
 DR Pfam; PF00024; PAN; 2.  
 DR Pfam; PF00100; Zona\_pellucida; 1.  
 DR SMART; SM00473; PAN\_AP; 3.  
 DR SMART; SM00241; ZP; 1.  
 DR PROSITE; PS50948; PAN; 3.  
 SQ SEQUENCE 774 AA; 83881 MW; 997A3B85350CECCC CRC64;  
 Query Match 9.2%; Score 125; DB 2; Length 774;  
 Best local similarity 25.6%; Pred. No. 0.0098;  
 Matches 62; Conservative 25; Mismatches 93; Indels 62; Gaps 13;  
 QY 21 CTSACLAADFTRGFCECSYMYVPVDAECIL-NTEDRLDRPLFVDEHEDYIYLDNNCA 78  
 DB 173 CLEAC-----QANSCSAVNY--ETGLCVMEFSTYDQLP-GSLSSQPVFTVYVQKSCF 224  
 QY 79 G---CECHMHFQNFRTSGIINDQCPAIAAQCYPVYQVAVAGEHQLSDHDSFEGLEL 135  
 DB 225 GVRPCKAKWICDR-----VQGRLEPRKAKSGSVATR 256  
 QY 136 SECCEICQRLSVLTANDFNCKSFMYSNLTRSCVLSDESRRLGRANLAIEVGMWTFESR 194  
 DB 257 RDCIELC-----LGEITFCRSANYVASGICELSDMRITLSDBANIAAYGGAUYLENN 311  
 QY 195 GV--PS-----FTRVQMLLVGFASFVMEVNSVTMCLDOQCSPPRETOGNFVCSVMY 248  
 DB 312 CAEPPSKLCEFRVAGRIKIVDS-VHQNVOYTLDECRLCLTAP-----FRCHS--YDY 362  
 QY 249 NE 250  
 DB 363 NE 364



RESULT 14  
Q94248 PRELIMINARY; PRT; 1127 AA.  
ID Q94248  
AC Q94248;  
DT 01-FEB-1997 (TREMBLrel. 02, Created)  
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)  
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)  
DE Hypothetical protein H03E18.1.  
GN Name=H03E18.1; ORFName=H03E18.1;  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematozoa; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderae; Caenorhabditis.  
NCBI\_TaxID=6229;  
RX MEDLINE=99069613; PubMed=9851916;  
RN SEQUENCE FROM N.A.  
RP STRAIN=Bristol N2;  
RC MEDLINE=99069613; PubMed=9851916;  
RG WormBase Consortium;  
RT "Genome sequence of the nematode C. elegans: a platform for  
investigating biology. The C. elegans Sequencing Consortium.";  
RL Science 282:2012-2018 (1998).  
[2]  
RA SEQUENCE FROM N.A.  
RB STRAIN=Bristol N2;  
RC Bentley D.;  
RT "The sequence of C. elegans fosmid H03E18."  
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.  
RN SEQUENCE FROM N.A.  
RP STRAIN=Bristol N2;  
RC Waterston R.;  
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
[4]  
RN SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RG WormBase Consortium;  
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U79477; AA007557.1; -.  
DR PIR; T25804; T25804.  
DR WormBase; WBGene00019148; H03E18.1.  
DR WormPep; H03E18.1; CB11558.  
DR InterPro; IPR003014; PAN.  
DR InterPro; IPR003609; Pan\_app.  
DR Pfam; PF00024; PAN; 5.  
DR SMART; SM00473; PAN AP; 5.  
DR PROSITE; PS00948; PAN; 4.  
KW Hypothetical protein.  
SEQUENCE 1127 AA; 124650 MW; E7A7608B10EAFPP2 CRC64;  
Query Match 9.2%; Score 124.5; DB 2; Length 1127;  
Best Local Similarity 22.3%; Pred. No. 0.017;  
Matches 44; Conservative 37; Mismatches 75; Indels 41; Gaps 9;  
1 EKKIVFARFVNSADSVRCASCLNAPDT--FGEECSVWYVVDACLTNEDRIDR 58  
Db 433 DDLHNVASVAGSEHDMNTEECQCMCKANSKTSRYEFCQASATYHAERDCILNEDRNK 492  
Qy 59 PDLFVDEHD-TVIVLDNNC-----AGCECHWHP-----NFKTSGIL 95  
Db 493 SKLFPEKQFTDVNVYITIGLACEVDETVKSGLALRECRKVVATGEPFKETSTKNGIX 552  
Qy 96 NDQQAFAIAQCYAPVYTVQVAVEGRQLSDDELHDSFEGLELSCBELCTQRLSVTANDPNC 155  
Db 553 SD-----DCYVE-LNDFV-LEGTALIVE-----TAVPEECKCKCAEGHLYGEE--C 596  
Qy 156 KSFMYSNLIRSCVLSDE 172  
Db 597 ASFLYYDSKTCLINKQ 613  
RESULT 15  
Q9VAG2

ID Q9VAG2 PRELIMINARY; PRT; 744 AA.  
AC Q9VAG2;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)  
DT CG7802-PA (CG7802-pb) (GH22837p).  
DE ORFName=CG7802;  
GN Drosophila melanogaster (Fruit fly).  
OS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
NCBI\_TaxID=7227;  
[1]  
RN SEQUENCE FROM N.A.  
RP MEDLINE=20196006; PubMed=1073113; DOI=10.1126/science.287.5461.2185;  
RX Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoekline R.A., Galley R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.H., Blazer V.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,  
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Baau A., Baxendale J., Bayraktaroglu L., Beasley B.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,  
RA Burtis K.C., Busam D.A., Butler H., Caudle E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glaeser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hoscin D., Houston K.A., Howland T.O., Wei M.H., Ibegyan C.,  
RA Jatali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M.,  
RA Palazolo M., Peltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun B.,  
RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodger, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
RA Yen R.P., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
RT "The genome sequence of Drosophila melanogaster.";  
RL Science 287:2185-2195 (2000).  
[2]  
RN SEQUENCE FROM N.A.  
RP MEDLINE=22426065; PubMed=12537568;  
RX Celniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,  
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Nelson A.,  
RA George R.A., Hoekline R.A., Laverly T., Muzny D.M., Nelson C.R.,  
RA Pacle J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
RA Svirkas R., Taber P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.,  
RT "Finishing a whole-genome shotgun Release 3 of the Drosophila  
melanogaster euchromatic genome sequence.";  
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079 (2002).  
[3]  
RN SEQUENCE FROM N.A.  
RP MEDLINE=22426070; PubMed=12537573;  
RX Kaminker J.S., Bergman C.M., Krommiller B., Carlson J., Svirkas R.,  
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
RA Ashburner M., Celniker S.E.,  
RT "The transposable elements of the Drosophila melanogaster euchromatin:  
a genomics perspective.";

